Algèbre 2
2016-2017

F. Sueur

Ce cours est inclus dans une combinaison convexe de

(1) cours d'algèbre linéaire 1 et 3 des années précédentes dues à C. Bachoc,

(2) cours d'algèbre linéaire 2 des années précédentes dues à J.-J. Ruch,

(3) lubies personnelles.

C'est avec gratitude que les deux premiers ingrédients sont employés.

Si vous trouvez des typos, erreurs, choses qui ne vont pas, merci de les signaler par mail à

franck.sueur@math.u-bordeaux.fr

Ce poly est donc amené à être actualisé de temps en temps.

Table des Matières

Chapitre 1

Le vocabulaire des lois de composition sur un ensemble

Ce premier chapitre est consacré à l'introduction de quelques structures algébriques fondamentales de l'algèbre générale, en particulier les groupes, les anneaux et les corps. Nous étudierons particulièrement le groupe symétrique qui sera très utile dans la suite du cours à travers notamment la notion de déterminant.

I. Groupes

Nous allons commencer par la notion de groupe. Le concept se développa à travers différents travaux notamment d' Évariste Galois, d'Augustin Louis Cauchy, d'Arthur Cayley, de Felix Klein, de Sophus Lie et c'est finalement Walther von Dyck pour la définition générale abstraite employée aujourd'hui.

I.1. Définition et premières propriétés. On commence par introduire la notion de groupe.

Définition 1. *Un groupe est un ensemble G muni d'une loi de composition interne $*$, c'est-à-dire d'une application :*

$$\begin{aligned} G \times G &\to G \\ (x, y) &\mapsto x * y \end{aligned}$$

vérifiant les propriétés suivantes :

(1) La loi $*$ est associative *: $x * (y * z) = (x * y) * z$ pour tout x, y, z dans G.*

(2) $(G, *)$ possède un élément neutre*, c'est-à-dire un élément $e \in G$ tel que $e * x = x * e = x$ pour tout x dans G.*

(3) Tout élément de G est inversible *: pour tout $x \in G$, il existe $y \in G$ tel que $x * y = y * x = e$.*

Si de plus, $x * y = y * x$ pour tout x, y appartenant à G, on dit que $(G, *)$ est un groupe commutatif (ou abélien, en l'honneur de Niels Abel).

Remarque 2. *On ne doit pas être impressionné par le vocabulaire employé ci-dessus : une loi de composition interne désigne seulement une opération interne dans cet ensemble, qui, à deux éléments quelconques de cet ensemble en associe, de manière unique, un troisième, toujours dans ce même ensemble.*

Remarque 3. *L'axiome d'associativité permet de définir l'opération sur trois éléments et non plus deux, en enlevant les parenthèses. Evidemment par une récurrence triviale le principe se généralise à un nombre quelconque fini d'éléments.*

Remarque 4. *Dans les exemples qui vont suivre, nous verrons que l'opération $*$ peut désigner suivant les cas l'addition $+$, la multiplication $\times$ ou encore la composition $\circ$. Il faut donc en pratique toujours avoir à l'esprit l'opération $*$ associée au groupe.*

Remarque 5. *Notons que l'on a précisé dans* (2) *que le neutre est neutre à gauche et à droite, et dans* (3) *que l'inverse est inverse à gauche et à droite car dans le cas général d'un groupe non commutatif le résultat des opérations $x * y$ et $y * x$ peuvent différer. Cependant les axiomes peuvent être réduits. Par exemple, on obtient une définition équivalente si on remplace les deux derniers axiomes de la définition ci-dessus par les conditions suivantes : il existe un élément e de G qui est neutre à gauche et tout élément admet un inverse à gauche, i.e. (si vous n'avez pas l'habitude, i.e. est une abréviation latine pour id est, ce qui veut dire c'est-à-dire) pour tout élément a de G, il existe b dans G tel que $b * a = e$ (ce qu'on peut exprimer en disant que tout élément de G admet un symétrique à gauche en association avec e).*

En particulier, si les conditions affaiblies sont vérifiées alors on peut montrer que les deux derniers axiomes de la définition sont vérifiés. En effet, supposons qu'il existe un élément e de G qui est neutre à gauche et tout élément admet un inverse à gauche et montrons successivement que tout élément admet aussi un inverse à droite puis que e est aussi élément neutre à droite.

Soit a un élément de G et b un inverse à gauche. Choisissons de même un inverse c à gauche de b. On a donc : $b*a=e$ et $c*b=e$. Donc $a*b=e*(a*b)$ car e est élément neutre à gauche. Ainsi $a*b=(c*b)*(a*b)$ car $c*b=e$ et donc $a*b=c*(b*a)*b$; par associativité. On obtient $a*b=c*e*b$; puisque $b*a=e$, puis $a*b=c*b$; car $e*b=b$, puisque e est élément neutre à gauche. Enfin $a*b=e$; car c est un inverse à gauche de b. Ainsi b est aussi un inverse à droite de a.

Prouvons maintenant que e est élément neutre à droite. Soit a un élément du groupe. D'après ce qui précède, nous pouvons choisir un élément b qui est inverse à gauche et à droite de a. Alors $a=e*a=(a*b)*a=a*(b*a)$ par associativité donc $a=a*e$ et donc e est bien aussi élément neutre à droite.

Proposition 6. *Dans un groupe $(G,*)$, on a les propriétés suivantes :*

(1) L'élément neutre est unique.

(2) L'inverse d'un élément est unique. On note x^{-1} l'inverse de $x\in G$.

(3) Le neutre vérifie $e^{-1}=e$.

(4) Pour tout x dans G, on a $(x^{-1})^{-1}=x$.

(5) Pour tout x et pour tout y dans G, on a $(x*y)^{-1}=y^{-1}*x^{-1}$.

(6) Pour tout a,x,y dans G, si $a*x=a*y$ ou si $x*a=y*a$, alors $x=y$.

Démonstration. On procède successivement.

(1) Si G possède deux neutres e et e' alors $e*e'=e$ mais aussi $e*e'=e'$ donc $e=e'$.

(2) Supposons que x ait deux inverses y et y'. Alors $x*y=e$ et $x*y'=e$ par définition. On en déduit que $x*y=x*y'$. On multiplie cette identité à gauche par y et on utilise l'associativité : $y*(x*y)=y*(x*y')$ donc $(y*x)*y=(y*x)*y'$. Mais $y*x=e$ donc $e*y=e*y'$ donc $y=y'$.

(3) Puisque e est neutre, en particulier, $e*e=e$. Mais e a un unique inverse donc cet inverse est bien e.

(4) De même, la propriété $x*x^{-1}=x^{-1}*x=e$ montre que x est l'inverse de x^{-1}.

(5) Enfin en utilisant deux fois l'associativité de la loi, que y est l'inverse de y^{-1}, puis que x est l'inverse de x^{-1}, on obtient :

$$\begin{aligned}(x*y)*(y^{-1}*x^{-1}) &= \big((x*y)*y^{-1}*\big)x^{-1}\\ &= \big(x*(y*y^{-1})*\big)x^{-1}\\ &= \big(x*e\big)*x^{-1}\\ &= x*x^{-1}\\ &= e.\end{aligned}$$

et on en conclut donc que l'inverse de $x*y$ est $y^{-1}*x^{-1}$.

(6) Il suffit d'utiliser l'inverse de a.

□

I.2. Exemples. Débutons avec quelques exemples standards :

— $(\mathbb{Z},+)$ pour lequel $e=0$, l'inverse de x est $-x$. Il est commutatif.

— $(\mathbb{Q}^*,\times)$ pour lequel $e=1$, l'inverse de x est $1/x$. Il est commutatif.

— L'ensemble $GL(n,\mathbb{R})$ des matrices carrées inversibles de taille n à coefficients dans $\mathbb{R}$, muni du produit des matrices. Le neutre est la matrice identité, l'inverse est la matrice inverse.

— Le groupe symétrique d'ordre n S_n des bijections σ de $\{1,2,\dots,n\}$ dans lui-même. Nous allons revenir en détails sur S_n par la suite. Ses éléments σ sont appelés des permutations de $\{1,2,\dots,n\}$. L'ensemble S_n, muni de la composition des applications est un groupe dont l'élément neutre est l'identité, noté Id. On note parfois la composition plus simplement $\sigma_1\sigma_2$ au lieu de $\sigma_1\circ\sigma_2$.

— Venons-en à l'exemple très important du **groupe** $(\mathbb{Z}/nZ, +)$**.** On définit une opération d'addition dans $\mathbb{Z}/n\mathbb{Z}$ en posant :

$$(a \bmod n) + (b \bmod n) = (a + b) \bmod n.$$

Pour avoir un sens, il faut montrer que cette définition ne dépend pas du choix d'un représentant d'une classe de congruence modulo n. Autrement dit, si $a = a' \bmod n$ et $b = b' \bmod n$, il faut montrer que $(a + b) = (a' + b') \bmod n$. Pour cela, on traduit les congruences par des égalités dans $\mathbb{Z}$: il existe u tel que $a = a' + un$ et il existe v tel que $b = b' + vn$. Alors $(a+b) = (a' + b') + (u+v)n$ ce qui conduit à : $(a + b) = (a' + b') \bmod n$. On vérifie aisément que le neutre pour cette opération est $0 \bmod n$ et que tout élément $x \bmod n$ est inversible, d'inverse $-x \bmod n$.

— On peut également utiliser **le groupe multiplicatif** $(\mathbb{Z}/n\mathbb{Z})^*$**.** De la même façon que pour l'addition, on peut définir une multiplication dans $\mathbb{Z}/n\mathbb{Z}$ en posant :

$$(a \bmod n)(b \bmod n) = (ab) \bmod n.$$

Cette loi est associative, commutative, et possède un élément neutre qui est $1 \bmod n$. Par contre, tout élément n'est pas inversible, en particulier $0 \bmod n$ n'est *jamais* inversible. Par exemple, on vérifie facilement que les inversibles modulo 4 sont 1 et 3.

Définition 7. *On note $(\mathbb{Z}/n\mathbb{Z})^*$ l'ensemble des éléments de $\mathbb{Z}/n\mathbb{Z}$ qui sont inversibles pour la multiplication.*

On a donc :

$$(\mathbb{Z}/n\mathbb{Z})^* = \{a \bmod n : \text{ il existe } b \in \mathbb{Z} \text{ tel que } ab = 1 \bmod n\}.$$

Alors $(\mathbb{Z}/n\mathbb{Z})^*$ muni de la multiplication est un groupe commutatif.

Par exemple :

$$(\mathbb{Z}/4\mathbb{Z})^* = \{1, 3\}, \quad (\mathbb{Z}/5\mathbb{Z})^* = \{1, 2, 3, 4\}, \quad (\mathbb{Z}/6\mathbb{Z})^* = \{1, 5\}.$$

En général, $(\mathbb{Z}/n\mathbb{Z})^*$ n'est donc pas $\mathbb{Z}/n\mathbb{Z}$ privé de 0.

Exercice 8. *Démontrez en détail que la multiplication définie ci-dessus a bien un sens, et que $(\mathbb{Z}/n\mathbb{Z})^*$ muni de cette multiplication est un groupe commutatif.*

Notations. Cette duplicité de structures de groupe lié à $\mathbb{Z}/n\mathbb{Z}$ nous amène aux quelques mots d'avertissement suivants. Pour alléger les notations, on va souvent utiliser la *notation multiplicative* pour un groupe général : $x * y = xy$, et $e = 1$. On dit que xy est le *produit* de x et y. On utilise aussi les raccourcis $x^n = x * \cdots * x$ (n termes), $x^0 = e$, pour $n \in \mathbb{N}$ et $x^{-n} = x^{-1} * \cdots * x^{-1}$.

Lorsque le groupe est commutatif, en particulier lorsque la loi est issue de l'addition usuelle, on emploie la *notation additive* $x * y = x + y$ et $e = 0$. Alors on parle d'opposé plutôt que d'inverse d'un élément, et on note $nx = x + \cdots + x$.

I.3. Sous-groupes. Introduisons dès maintenant la notion de sous-groupes.

Définition 9. *Soit $(G, *)$ un groupe. Un sous-groupe de G est un sous-ensemble $H \subset G$ tel que $(H, *)$ soit un groupe.*

Nous allons donner deux caractérisations de la notion de sous-groupes et donnerons des exemples seulement ensuite car ce sera plus simple de montrer que ce sont des sous-groupes avec la deuxième caractérisation. Enfin nous verrons comment se comporte cette notion de sous-groupes par produit cartésien, intersection et union.

Proposition 10. *Soit $(G, *)$ un groupe et soit $H \subset G$. Alors H est un sous-groupe de G si et seulement si*

(i) *Pour tout $x \in H, y \in H, x * y \in H$.*
(ii) $e \in H$
(iii) *Pour tout $x \in H, x^{-1} \in H$.*

Démonstration. Examinons les propriétés que doit vérifier $H \subset G$ pour être un groupe pour $*$.

Tout d'abord il est nécessaire que la loi $*$ soit interne dans H, c'est-à-dire que $x * y \in H$ pour tout $x \in H$, $y \in H$. Remarquons que la loi $*$ étant associative dans G, elle l'est forcément dans H.

Le neutre de H ne peut être que le neutre de G. En effet, si e' est le neutre de H, on a comme $e' \in G$, $e' * e = e'$. Mais aussi $e' * e' = e'$ en raisonnant dans H. Donc $e' * e = e' * e'$. Comme $e' \in G$, il a un inverse e'^{-1} dans G. On multiplie la précédente égalité à gauche par celui-ci, pour obtenir : $(e'^{-1} * e') * e = (e'^{-1} * e') * e'$ soit $e * e = e * e'$ soit $e = e'$. Donc si $(H, *)$ est un groupe, son neutre est e le neutre de G.

Pour ce qui est de l'inverse, un élément de H a bien toujours un inverse (unique) dans G. Il faut donc que cet inverse appartienne à H.

Réciproquement, si les propriétés (i), (ii), (iii), sont vérifiées, alors $(H, *)$ est bien un groupe. En effet, l'associativité et la propriété $e * x = x$ sont automatiquement vraies dans H puisqu'elles sont vraies dans G. □

La proposition suivante énonce la propriété minimale suffisante à vérifier pour qu'un sous-ensemble de G soit un sous-groupe de G.

Proposition 11. *Soit $(G, *)$ un groupe et soit $H \subset G$. Alors H est un sous-groupe de G si et seulement si, H est non vide, et vérifie :*

$$\text{Pour tout } x \in H, y \in H, x * y^{-1} \in H. \tag{1}$$

Démonstration. Supposons que H soit un sous-groupe de G. Alors on a vu que H vérifie (i), (ii), (iii). En particulier (ii) indique que le neutre e de G appartient à H donc H est non vide. Si x et y sont dans H, d'après (iii), $y^{-1} \in H$ et, d'après (i), $x * y^{-1} \in H$.

Réciproquement, supposons que H est non vide et vérifie (1). Il contient donc un élément x. Donc, par (1), $x * x^{-1} = e \in H$. En appliquant encore (1), on obtient $e * x^{-1} = x^{-1} \in H$. Si x et y sont dans H, on a vu que $y^{-1} \in H$, donc encore avec (1), $x * y = x * (y^{-1})^{-1} \in H$. Donc (i), (ii), (iii) sont vérifiées donc H est bien un sous-groupe de G. □

Exemple 12.

- *$\{e\}$ et G sont des sous-groupes de G.*
- *Pour tout n entier naturel, l'ensemble $n\mathbb{Z} := \{nq,\, q \in \mathbb{Z}\}$ est un sous-groupe de $(\mathbb{Z}, +)$. En effet, il est non vide puisque $0 \in n\mathbb{Z}$ et si $x = nk \in n\mathbb{Z}$, $y = n\ell \in \mathbb{Z}$, $x - y = n(k - \ell) \in n\mathbb{Z}$.*

Proposition 13. *Les sous-groupes de $(\mathbb{Z}, +)$ sont les $n\mathbb{Z}$, pour n dans $\mathbb{N}$.*

Démonstration. On vient de voir que $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$. Réciproquement, soit H un sous-groupe de $\mathbb{Z}$. Soit n son plus petit élément strictement positif. Soit $x \in H$, par division euclidienne il existe q dans $\mathbb{Z}$ et $r \in \{0, 1, \ldots, n-1\}$ tels que $x = qn + r$. Alors qn, et donc $r = x - qn$ appartient à H. Comme $r \in \{0, 1, \ldots, n-1\} \cap H$ et que n est le plus petit élément strictement positif de H, il n'y a qu'une possibilité c'est $r = 0$. Donc $x \in n\mathbb{Z}$ et $H = n\mathbb{Z}$. □

Evoquons maintenant la stabilité de la notion de sous-groupes vis-à-vis des opérations ensemblistes usuelles.

Notons d'abord que si l'on dispose de deux groupes $(G, *)$ et $(G', \circ)$, on peut définir le *produit direct* $G \times G'$ de ces deux groupes comme l'ensemble

$$G \times G' = \{(x, y) : x \in G, y \in G'\}.$$

Muni de l'opération : $(x, y) \cdot (x', y') = (x * x', y \circ y')$, c'est un groupe dont le neutre est $(e_G, e_{G'})$ et l'inverse de (x, y) est (x^{-1}, y^{-1}).

La proposition suivante, relative à l'intersection, dont la preuve est laissée au lecteur, sera utile pour la suite.

Proposition 14. *Soit H_1 et H_2 deux sous-groupes de $(G, *)$. L'intersection $H_1 \cap H_2$ de H_1 et H_2 est un sous-groupe de G.*

Attention, en revanche la réunion de deux sous-groupes n'est pas un sous-groupe. Par exemple $2\mathbb{Z} \cup 3\mathbb{Z}$ n'est pas un sous-groupe de $(\mathbb{Z}, +)$ car $2 + 3 = 5$ n'est pas dans cet ensemble.

I.4. Sous-groupes engendrés par une partie. On en vient à la notion de sous-groupe engendré par une partie.

Définition 15. *Soit $S \subset G$. On appelle* sous-groupe engendré par S *et on note $\langle S \rangle$ l'intersection de tous les sous-groupes de G contenant S.*

Il est facile de vérifier que c'est un sous-groupe de G, et que c'est le plus petit contenant S (au sens où, si H est un sous-groupe de G contenant S, alors $\langle S \rangle \subset H$).

Si $S = \{x\}$ avec $x \in G$, on note $\langle x \rangle = \langle \{x\} \rangle$.

Exemple 16.

(1) Si $S = \{e\}$, $\langle e \rangle = \{e\}$

(2) Si $S = \{2, 3\} \subset \mathbb{Z}$, $\langle S \rangle = \mathbb{Z}$. En effet, $1 = 3 - 2 \in \langle S \rangle$.

Proposition 17. *Soit G un groupe noté multiplicativement.*

(1) $\langle x \rangle = \{x^k : k \in \mathbb{Z}\}$.

(2) Si x et y commutent, i.e. $xy = yx$, alors $\langle x, y \rangle = \{x^k y^\ell : k, \ell \in \mathbb{Z}\}$.

(3) Si H_1 et H_2 sont des sous-groupes de G, et que $h_1 h_2 = h_2 h_1$ pour tout $h_1 \in H_1$, $h_2 \in H_2$, alors $\langle H_1 \cup H_2 \rangle = \{h_1 h_2 : h_1 \in H_1, h_2 \in H_2\}$.

La preuve est laissée au lecteur.

À partir de maintenant, on utilise systématiquement la notation multiplicative $x * y = xy$ et $e = 1_G$ pour un groupe général G. Un peu plus tard on simplifiera encore 1_G en 1.

I.5. Ordres. Définissons successivement les notions d'ordre d'un groupe et d'ordre d'un élément.

Définition 18. *L'ordre d'un groupe est le nombre de ses éléments. On note $|G|$ l'ordre de G.*

Définition 19. *Soit G un groupe et soit $x \in G$. L'ordre de x est le plus petit entier $k \geq 1$, s'il existe, tel que $x^k = 1_G$.*

Si pour tout k, $x^k \neq 1_G$, on dit que x est d'ordre infini.

Exemple 20.

- *Le neutre 1_G est toujours d'ordre 1.*
- *Dans $(\mathbb{Z}, +)$, tout élément non nul est d'ordre infini.*
- *Dans $(\mathbb{Z}/n\mathbb{Z}, +)$, tout élément a vérifie $na = 0$. Mais a n'est pas forcément d'ordre n ! Exemple : dans $\mathbb{Z}/6\mathbb{Z}$, 2 est d'ordre 3.*
- *Dans $(\mathbb{Z}/n\mathbb{Z}, +)$, 1 est d'ordre n.*
- *Dans S_n, un cycle de longueur p est d'ordre p. En particulier, les transpositions sont d'ordre 2.*

Proposition 21. *Soit G un groupe et soit $x \in G$, un élément d'ordre k. Si $n \in \mathbb{Z}$ et si $n = kq + r$, avec q, r dans $\mathbb{Z}$ et $0 \leq r < k$ est la division euclidienne de n par k, alors*

$$x^n = x^r.$$

On a l'équivalence :

$$x^n = 1_G \iff k \text{ divise } n$$

Démonstration. Si $n = kq + r$ alors $x^n = x^{kq+r} = (x^k)^q . x^r$. Donc, si $x^k = 1$ alors $x^n = x^r$.

En particulier, si k divise n alors $r = 0$ et $x^n = x^r = 1$. Réciproquement, si $x^n = 1$, alors $x^r = 1$ avec $0 \leq r < k$ mais comme k est le plus petit entier positif avec cette propriété, c'est que $r = 0$ et donc que k divise n. □

Remarque 22.

(1) *Si $x^n = 1_G$, il ne faut pas conclure trop rapidement que n est l'ordre de x. Par contre, on sait que l'ordre de x est un diviseur de n, ça limite les possibilités. En fait, on peut alors calculer l'ordre de x en descendant* l'arbre des diviseurs de n.

(2) *Si G est un groupe fini, alors tout élément est d'ordre fini. En effet, $\{1_G, x, x^2, \ldots, x^n, \ldots\}$ ne peut être infini donc il existe $k < \ell$ tels que $x^k = x^\ell$ d'où on tire $x^{\ell-k} = 1_G$.*

On a déjà vu la notion de sous-groupe engendré par un élément $x \in G$. Cette notion va justement nous permettre de définir la notion de groupe cyclique.

Définition 23. *Un groupe G est dit* monogène *s'il existe $x \in G$ tel que $G = \langle x \rangle$. Un groupe monogène fini est dit* cyclique. *Un élément x tel que $G = \langle x \rangle$ est appelé* un générateur *de G.*

Exemple 24. *$\mathbb{Z}/n\mathbb{Z}$ est cyclique d'ordre n. $(\mathbb{Z}/5\mathbb{Z})^*$ est cyclique d'ordre 4. Listez leurs générateurs.*

Proposition 25. *Soit G un groupe et $x \in G$.*

(1) *Si x est d'ordre k, $\langle x \rangle = \{1, x, x^2, \ldots, x^{k-1}\}$ et $|\langle x \rangle| = k$.*

(2) *G est cyclique si et seulement si G contient un élément d'ordre $|G|$.*

Démonstration. On a déjà vu que $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$ donc $\{1, x, x^2, \ldots, x^{k-1}\} \subset \langle x \rangle$. Si x est d'ordre k, $x^n = x^r$ où r est le reste de n dans la division euclidienne par k, $0 \leq r < k$, donc l'inclusion inverse est vérifiée. Il reste à montrer que l'ensemble $\{1, x, x^2, \ldots, x^{k-1}\}$ a exactement k éléments, c'est-à-dire que les x^i, $0 \leq i \leq k-1$ sont distincts. En effet, supposons que, pour $0 \leq i < j \leq k-1$, on ait $x^i = x^j$. Alors, $x^{j-i} = 1$. Mais $1 \leq j - i \leq k-1$, donc c'est en contradiction avec la propriété que k est le plus petit entier positif tel que $x^k = 1$.

Supposons G cyclique. Alors, il existe $x \in G$ tel que $G = \langle x \rangle$, et, d'après la discussion qui précède, $|G|$ est égal à l'ordre de x. Donc G contient bien un élément dont l'ordre vaut $|G|$. Réciproquement, supposons que G contienne un élément x d'ordre $k = |G|$. Alors $\langle x \rangle \subset G$ et $|\langle x \rangle| = k = |G|$ donc on peut conclure que $\langle x \rangle = G$ □

Attention : Un groupe cyclique n'a pas un unique générateur. En fait, il a autant de générateurs qu'il y a d'éléments d'ordre égal à l'ordre de ce groupe.

II. Le groupe symétrique S_n

II.1. Définition. On commence par rappeler la notion de permutation, ce seront les éléments du groupe symétrique.

Définition 26. *Une* permutations σ *de $\{1, 2, \ldots, n\}$ est une bijection de $\{1, 2, \ldots, n\}$ dans lui-même.*

Définition 27. *On appelle groupe symétrique d'ordre n et on note S_n l'ensemble des*permutations *de $\{1, 2, \ldots, n\}$ muni de la loi de composition.*

Une permutation est notée de la façon suivante :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}.$$

Un exemple crucial de permutation est celui des transpositions.

Définition 28. *Une* transposition τ *est une permutation qui laisse invariant tous les éléments sauf deux qu'elle échange : il existe deux éléments distincts i et j tels que :*

$$\tau(i) = j \text{ et } \tau(j) = i \text{ et } \quad \forall k \neq i \text{ et } k \neq j,\ \tau(k) = k.$$

Voyons maintenant une autre classe de permutations qui généralise la notion de transposition.

Définition 29. *Un* cycle *est une permutation particulière qui permute circulairement un sous-ensemble de $\{1, \ldots, n\}$ et laisse les autres éléments inchangés. On note le cycle $\sigma = (a_1, \ldots, a_p)$, $p \geq 2$, si $\sigma(a_1) = a_2$, $\sigma(a_2) = a_3, \ldots, \sigma(a_p) = a_1$. On dit que $p \geq 2$ est* la longueur *du cycle σ et on la note $\ell(\sigma)$. Une* transposition *est un cycle de longueur* 2.

Une transposition est donc un cycle de longueur 2.

Exemple 30. *Un cycle de longueur* 3 *dans* S_5.

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix} = (1,3,2).$$

Exemple 31. *Les éléments de* S_3 *sont*

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad (1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$(2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad (1\ 2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \text{ et } (1\ 3\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On remarque qu'il y a $6 = 3!$ *éléments.*

Exercice 32. *Montrez que* S_n *n'est pas commutatif si* $n \geq 3$.

Exercice 33. *Montrez que* S_n *a* $n!$ *éléments c'est-à-dire qu'il est d'ordre* $n!$.

Définition 34. *Le* support *d'une permutation* σ *est l'ensemble :*

$$\mathrm{Sup}(\sigma) := \{i \in \{1,\dots,n\} : \sigma(i) \neq i\}.$$

Proposition 35. *On a les propriétés suivantes :*

(1) Si $\sigma = (a_1,\dots,a_p)$, $\mathrm{Sup}(\sigma) = \{a_1,\dots,a_p\}$.

(2) Si S *est le support de* σ, $\sigma(S) = S$.

(3) Deux permutations de supports disjoints commutent.

Démonstration. 1. et 2. sont faciles. Montrons 3. Soit σ_1 une permutation de support S_1 et σ_2 une permutation de support S_2, telles que $S_1 \cap S_2 = \emptyset$. Montrons que $\sigma_1\sigma_2 = \sigma_2\sigma_1$. On peut distinguer trois cas :

— $i \notin S_1 \cup S_2$. Alors, $\sigma_1\sigma_2(i) = \sigma_1(\sigma_2(i)) = \sigma_1(i) = i$, et de même, $\sigma_2\sigma_1(i) = \sigma_2(\sigma_1(i)) = \sigma_2(i) = i$.

— $i \in S_1$. Alors, $i \notin S_2$ donc $\sigma_1\sigma_2(i) = \sigma_1(i)$. De plus, $\sigma_1(i) \in S_1$ d'après 2. donc $\sigma_1(i) \notin S_2$ et $\sigma_2(\sigma_1(i)) = \sigma_1(i)$.

— $i \in S_2$. Ce cas est analogue au précédent.

□

Proposition 36. *Un cycle de longueur* p *est d'ordre* p *dans* S_n.

Démonstration. Si $c = (a_1,\dots,a_p)$, $c^k(a_1) = a_{k+1}$ donc, si $c^k = 1$, alors $k \geq p$. Il est clair que $c^p = 1$. □

II.2. La décomposition canonique d'une permutation en produit de cycles.

Nous admettrons le résultat suivant.

Théorème 37.

Une permutation σ *se décompose en un produit de cycles à supports disjoints, de façon unique à l'ordre près.*

Voyons cela sur un exemple :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 3 & 1 & 7 & 6 & 5 & 2 & 8 & 10 & 9 \end{pmatrix}$$

On part de 1 et on applique successivement σ pour obtenir le premier cycle $c_1 : 1 \to 4 \to 7 \to 2 \to 3 \to 1$, donc $c_1 = (1, 4, 7, 2, 3)$. On recommence en partant d'un élément qui n'a pas été visité : $5 \to 6 \to 5$; puis $8 \to 8$ et $9 \to 10 \to 9$ qui donnent :

$$\sigma = (1, 4, 7, 2, 3)(5, 6)(9, 10).$$

Observons que ceci peut être très utile pour calculer les itérés. Prenons par exemple pour objectif de calculer la permutation σ^{1145}. Parce que les cycles à supports disjoints commutent,

$$\sigma^{1145} = (1, 4, 7, 2, 3)^{1145}(5, 6)^{1145}(9, 10)^{1145}.$$

Il suffit maintenant de réduire les exposants modulo les longueurs respectives des cycles, ce qui donne

$$\sigma^{1145} = (1, 4, 7, 2, 3)^0(5, 6)^1(9, 10)^1 = (5, 6)(9, 10).$$

Corollaire 38. *Soit $\sigma = c_1 \dots c_s$ la décomposition en produits de cycles disjoints de σ, avec $\ell_i := \ell(c_i)$ ordonnés par ordre décroissant. L'ordre de σ est le ppcm de $\ell_1, \dots, \ell_s$.*

Démonstration. On a déjà vu que $\sigma^k = c_1^k \dots c_s^k$. Comme les supports des c_i sont disjoints, $\sigma^k = 1$ si et seulement si $c_i^k = 1$ pour tout $i = 1, \dots, s$. Comme l'ordre de c_i vaut ℓ_i, $c_i^k = 1$ si et seulement si ℓ_i divise k. Finalement, on a démontré que $\sigma^k = 1$ si et seulement si $\mathrm{ppcm}(\ell_1, \dots, \ell_s)$ divise k donc l'ordre de σ est bien $\mathrm{ppcm}(\ell_1, \dots, \ell_s)$. □

Examinons maintenant d'autres décompositions des permutations.

Théorème 39.

On a :

(1) $c = (a_1, \dots, a_p) = (a_1, a_2)(a_2, a_3) \dots (a_{p-1}, a_p)$.

(2) Toute permutation est un produit de transpositions.

Démonstration. Le premier point se vérifie directement, et le deuxième en résulte en combinant avec le théorème 37. □

Corollaire 40. *Les ensembles suivants sont générateurs de S_n :*

(1) L'ensemble des transpositions

(2) $\{(1, 2), (1, 3), \dots, (1, k), \dots, (1, n)\}$

(3) $\{(1, 2), (2, 3), \dots, (k - 1, k), \dots, (n - 1, n)\}$

(4) $\{(1, 2), (1, 2, 3, \dots, n)\}$

Démonstration. (1) C'est le théorème 39.

(2) On a $(i, j) = (1, i)(1, j)(1, i)$.

(3) On a $(1, i + 1) = (1, i)(i, i + 1)(1, i)$.

(4) Soit $c = (1, 2, \dots)$; $c(1, 2)c^{-1} = (2, 3)$, etc..

□

II.3. La signature. On commence par définir le nombre d'inversions d'une permutation.

Définition 41. *Soit $\sigma \in S_n$. Le* nombre d'inversions de σ, *est le nombre*

$$\mathcal{I}(\sigma) := |\{(i,j) : 1 \leq i < j \leq n \text{ et } \sigma(i) > \sigma(j)\}|.$$

Nous rappelons que les barres verticales ci-dessus désigne le cardinal de l'ensemble.
Ceci nous permet de définir la notion de signature d'une permutation.

Définition 42. *La* signature de σ *est définie par :*

$$\epsilon(\sigma) = (-1)^{\mathcal{I}(\sigma)}.$$

Une permutation est dite paire quand elle présente un nombre pair d'inversions, impaire sinon. La signature d'une permutation paire est 1 ; celle d'une permutation impaire est ?1.

Exemple 43. *Soit la permutation*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}.$$

La liste des paires en inversion est {2, 5}, {3, 4}, {3, 5}, {4, 5}. Il y en a quatre, donc la signature est 1 *et la permutation est paire.*

Proposition 44. *Une transposition est une permutation impaire.*

Démonstration. Notons $i < j$ les termes échangés par la transposition, de sorte que

$$\begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}.$$

Les paires en inversion sont les paires de la forme $\{i,k\}$ avec k compris entre $i+1$ et j et celles de la forme $\{k,j\}$ avec k compris entre $i+1$ et j?1. Au total, il y a $(j-i)+(j-i+1)$ soit un nombre impair d'inversions, et l'imparité de la permutation en découle. □

On la reformulation suivante que l'on admettra.

Proposition 45. *On a*

$$\epsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Cette expression n'est pas très utile en pratique mais elle permet de montrer facilement le résultat suivant sur la multiplicativité de la signature.

Théorème 46.

Pour toutes permutations $\sigma, \sigma' \in S_n$,

$$\epsilon(\sigma\sigma') = \epsilon(\sigma)\epsilon(\sigma').$$

Démonstration. On calcule $\epsilon(\sigma\sigma')$ avec la formule de la proposition 45 :

$$\begin{aligned} \epsilon(\sigma\sigma') &= \prod_{1 \leq i < j \leq n} \frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{i - j} \\ &= \prod_{1 \leq i < j \leq n} \Big(\frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{\sigma'(i) - \sigma'(j)}\Big)\Big(\frac{\sigma'(i) - \sigma'(j)}{i - j}\Big) \\ &= \prod_{1 \leq i < j \leq n} \frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{\sigma'(i) - \sigma'(j)} \prod_{1 \leq i < j \leq n} \frac{\sigma'(i) - \sigma'(j)}{i - j}. \end{aligned}$$

On remarque que, en posant $i' = \sigma'(i)$ et $j' = \sigma'(j)$, on a

$$\frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{\sigma'(i) - \sigma'(j)} = \frac{\sigma(i') - \sigma(j')}{i' - j'},$$

donc

$$\epsilon(\sigma\sigma') = \prod_{1\leq i'<j'\leq n} \frac{\sigma(i') - \sigma(j')}{i' - j'} \prod_{1\leq i<j\leq n} \frac{\sigma'(i) - \sigma'(j)}{i - j} = \epsilon(\sigma)\epsilon(\sigma').$$

□

Nous allons énoncer deux corollaires de ce théorème. Tout d'abord, nous examinons la signature des cycles.

Corollaire 47. *Si c est un cycle de longueur ℓ, $\epsilon(c) = (-1)^{\ell-1}$.*

Démonstration. On a vu que $\epsilon((i,j)) = -1$ et on combine le premier point du théorème 39 avec le résultat précédent. □

Enfin nous laisserons le soin au lecteur de déduire le résultat suivant du résultat 46.

Corollaire 48. *L'ensemble des permutations de signature $+1$ est un sous-groupe de S_n, appelé* le groupe alterné *et noté A_n.*

II.4. Complément : conjugaison dans S_n. En guise de complément voyons comment se comporte le groupe S_n sous l'effet des conjugaisons.

Proposition 49.

(1) Si $c = (a_1, \ldots, a_p)$, et $\sigma \in S_n$, alors $\sigma c \sigma^{-1}$ est encore un cycle, de même longueur que c :

$$\sigma c \sigma^{-1} = (\sigma(a_1), \ldots, \sigma(a_p)).$$

(2) Deux cycles de même longueur sont conjugués dans S_n.

(3) Soit $\sigma = c_1 \ldots c_s$ la décomposition en produits de cycles disjoints de σ, avec $\ell_i := \ell(c_i)$ ordonnés par ordre décroissant. Alors deux permutations sont conjuguées dans S_n si et seulement si elles ont le même $(\ell_1, \ldots, \ell_s)$.

Démonstration. Nous allons procéder point par point.

(1) Posons $\tau = \sigma c \sigma^{-1}$. Alors, $\tau(\sigma(a_i)) = \sigma c \sigma^{-1}\sigma(a_i) = \sigma c(a_i) = \sigma(a_{i+1})$. D'autre part, si $k \notin \mathrm{Sup}(c)$,

$$\tau(\sigma(k)) = \sigma c \sigma^{-1}\sigma(k) = \sigma c(k) = \sigma(k).$$

Donc τ est bien le cycle $(\sigma(a_1), \ldots, \sigma(a_p))$.

(2) Si $c = (a_1, \ldots, a_p)$, et $c' = (b_1, \ldots, b_p)$, il existe une permutation σ de $\{1, 2, \ldots, n\}$ telle que $\sigma(a_i) = b_i$ pour $i = 1, \ldots, p$. D'après le deuxième point $c' = \sigma c \sigma^{-1}$.

(3) Soit σ et σ' deux permutations conjuguées dans S_n. Il existe donc $\tau \in S_n$ tel que $\sigma' = \tau\sigma\tau^{-1}$. Avec les notations du théorème, on en déduit que

$$\begin{aligned}\sigma' = \tau\sigma\tau^{-1} &= \tau c_1 \ldots c_s \tau^{-1} \\ &= (\tau c_1 \tau^{-1})(\tau c_2 \tau^{-1}) \ldots (\tau c_s \tau^{-1}).\end{aligned}$$

D'après ce qui précède, $c'_i := \tau c_i \tau^{-1}$ est un cycle de même longueur que c_i et de support $\tau(\mathrm{Sup}(c_i))$. Donc $\sigma' = c'_1 \ldots c'_s$ est son unique décomposition en produit de cycles disjoints. Si $\ell'_i := \ell(c'_i)$, on a donc $(\ell_1, \ldots, \ell_s) = (\ell'_1, \ldots, \ell'_s)$.

Réciproquement, supposons que σ et σ' soient deux permutations dont les decompositions en produit de cycles disjoints

$$\sigma = \prod_{i-1}^{s} c_i \quad \text{et} \quad \sigma' = \prod_{i=1}^{s} c'_i$$

vérifient $\ell(c_i) = \ell(c'_i) =: \ell_i$. Alors, on peut construire une permutation τ telle que

$$c_i = (a_{i,1}, \ldots, a_{i,\ell_i}), \quad c'_i = (b_{i,1}, \ldots, b_{i,ell_i}), \text{ et } \tau(a_{i,j}) = b_{i,j}.$$

On vérifie que $\sigma' = \tau\sigma\tau^{-1}$. Donc σ et σ' sont conjuguées.

□

III. Anneaux et corps

III.1. Définitions. On commence par la définition de la notion assez précieuse d'anneaux.

Définition 50. *Un* anneau $(A, +, \cdot)$ *est un ensemble muni de deux lois de composition interne $+$ et $\cdot$ telles que :*

(1) $(A, +)$ *est un groupe commutatif de neutre noté* 0 *et appelé* zéro.

(2) La loi $\cdot$

(a) est associative : $(x \cdot y) \cdot x = x \cdot (y \cdot z)$ *pour tout* $x, y, z \in A$.

(b) possède un élément neutre noté 1 *et appelé* un *ou* unité *:* $1 \cdot x = x \cdot 1 = x$ *pour tout* $x \in A$.

(3) La loi $\cdot$ *est distributive sur l'addition* $+$ *: pour tout* $x, y, z \in A$, $x \cdot (y + z) = x \cdot y + x \cdot z$ *et* $(x + y) \cdot z = x \cdot z + y \cdot z$.

Si en outre la loi $\cdot$ *est commutative, on dit que* A *est un* anneau commutatif.

Remarque 51. *On a :* $0 \cdot x = x \cdot 0 = 0$ *pour tout* $x \in A$. *En effet,*

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x \Longrightarrow 0 \cdot x = 0.$$

Également, on montre que $(-x) \cdot y = -(x \cdot y)$ *car :*

$$x \cdot y + (-x) \cdot y = (x + (-x)) \cdot y = 0 \cdot y = 0.$$

L'opposé de x pour $+$ est noté $-x$. Désormais on note $x \cdot y = xy$.

Définition 52. *Soit* $(A, +, \cdot)$ *un anneau. Un élément* $x \in A$ *est appelé* un inversible *(ou une unité) de* A *s'il est inversible pour* $\cdot$, *c'est-à-dire s'il existe* $y \in A$ *tel que* $xy = yx = 1$. *On note* $y = x^{-1}$. *L'ensemble des inversibles de* A *est noté* A^*.

Proposition 53. $(A^*, \cdot)$ *est un groupe appelé le groupe des inversibles (ou le groupe des unités) de* A. *On a* $A^* \subset A \setminus \{0\}$. *Si* A *est commutatif et si* $A^* = A \setminus \{0\}$, *c'est-à-dire si tout élément non nul de* A *est inversible, on dit que* A est un corps.

Démonstration. La loi $\cdot$ est bien interne dans A^* : si $x, y \in A^*$ alors xy est inversible d'inverse $y^{-1}x^{-1}$. On a $1 \in A^*$ car 1 est inversible : $1 \cdot 1 = 1$. Par définition, tout élément $x \in A^*$ est inversible, et son inverse x^{-1} est aussi dans A^* puisque il est inversible d'inverse x.

On a vu que $0x = 0$ donc 0 n'est jamais inversible, d'où l'inclusion $A^* \subset A \setminus \{0\}$. □

Définition 54. *Un* diviseur de zéro *d'un anneau* A *est un élément* $x \in A$ *tel que :* $x \neq 0$ *et il existe* $y \neq 0$, $y \in A$, *avec* $xy = 0$ *ou* $yx = 0$. *Un anneau* A *sans diviseur de zéro s'appelle un* anneau intègre.

Remarque 55. *Si* $x \in A^*$ *alors* x *n'est pas un diviseur de zéro de* A. *En effet, si* $xy = 0$, *en multipliant à gauche par* x^{-1}, *on obtient* $y = 0$.

III.2. Exemples. On commence avec par remarquer que les ensembles les plus courants ont une structure d'anneaux.

Exemple 56. *Les ensembles* $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ *sont des anneaux commutatifs et intègres, pour les opérations d'addition et de multiplication usuelles.* $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ *sont des corps.*

Exemple 57. $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ *est un anneau. Son groupe des unités est* $(\mathbb{Z}/n\mathbb{Z})^*$*, d'ordre* $\varphi(n)$*.* $\mathbb{Z}/n\mathbb{Z}$ *est un corps si et seulement si* n *est premier. Si* n *n'est pas premier, les éléments non nuls et non inversibles, c'est-à-dire les* $a \bmod n$ *tels que* $\mathrm{pgcd}(a, n) > 1$*, sont tous diviseurs de zéro.*

Exemple 58. *Si* A *est un anneau commutatif, l'ensemble* $M_n(A)$ *des matrices carrées de taille* n *est un anneau de zéro la matrice dont tous les coefficients font* 0 *et de* 1 *la matrice identité* Id_n*. Si* $n \geq 2$ *il n'est pas commutatif et possède des diviseurs de zéros.*

Exemple 59. *Si* A *est un anneau commutatif, l'ensemble* $A[X]$ *des polynômes à coefficients dans* A *est un anneau commutatif.*

$$A[X] = \{P(X) = \sum_{k=0}^{n} a_k X^k : n \geq 0, (a_0, \ldots, a_n) \in A^{n+1}\}.$$

Les opérations d'addition et de multiplication sont définies, pour

$$P(X) = \sum_{k=0}^{n} a_k X^k \text{ et } Q(X) = \sum_{k=0}^{n} b_k X^k,$$

par :

$$(P + Q)(X) = \sum_{k=0}^{n} (a_k + b_k) X^k$$

et

$$PQ(X) = \sum_{k=0}^{2n} c_k X^k, \quad c_k = \sum_{\substack{0 \leq i \leq n \\ 0 \leq j \leq n \\ i+j=k}} a_i b_j.$$

Le zéro de $A[X]$ *est* $P = 0_A$*, le neutre pour la multiplication est* $P = 1_A$*.*

Exemple 60. *L'anneau des polynômes à* n *indéterminées* $A[X_1, \ldots, X_n]$ *et à coefficients dans un anneau commutatif* A *généralise l'exemple précédent. Il peut être construit récursivement :* $A[x_1, \ldots, X_n] = (A[X_1, \ldots, X_{n-1}])[X_n]$*.*

On résume les propriétés de ces anneaux dans le tableau suivant :

A	A^*	diviseurs de zéro	commutatif	intègre	corps
$\mathbb{Z}$	$\{-1, 1\}$	$\emptyset$	oui	oui	non
$K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$	$K \setminus \{0\}$	$\emptyset$	oui	oui	oui
$\mathbb{Z}/n\mathbb{Z}$	$\{a \bmod n : \mathrm{pgcd}(a, n) = 1\}$	$\{a \neq 0 \bmod n : \mathrm{pgcd}(a, n) > 1\}$	oui	ssi n premier	ssi n premier
$M_n(K), n \geq 2$ K corps	$GL_n(K) = \{M : \det(M) \neq 0\}$	$\{M \neq 0 : \det(M) = 0\}$	non	non	non
$K[X_1, \ldots, X_n]$ K corps	K^*	$\emptyset$	oui	oui	non

III.3. Sous-anneaux, idéaux, anneaux quotients. On traite ici du cas d'un anneau commutatif.

Définition 61. *Un sous-anneau B d'un anneau commutatif $(A, +, \cdot)$ est un sous-ensemble de A qui est un anneau pour les mêmes lois, et qui contient 1_A.*

Proposition 62. *$B \subset A$ est un sous-anneau de A s'il vérifie les conditions suivantes :*

(1) $1_A \in B$

(2) Pour tout $x, y \in B$, $x - y \in B$

(3) Pour tout $x, y \in B$, $xy \in B$

Démonstration. Les conditions 1. et 2. garantissent que $(B, +)$ est un sous-groupe de $(A, +)$ d'après la proposition 11, et la condition 3. que la mutiplication est une loi de composition interne pour B. Comme $1_A \in B$, la multiplication possède bien un élément neutre dans B. Les autres propriétés définissant un anneau sont vraies pour A donc aussi pour B. □

Définition 63. *Un sous-ensemble $I \subset A$ est un* idéal *de A si :*

(1) $(I, +)$ est un sous-groupe de $(A, +)$.

(2) Pour tout $x \in I$, et tout $a \in A$, $ax \in I$.

Exemple 64. *$\{0\}$ et A sont des idéaux de A.*

Proposition 65. *Soit I un idéal de A. Alors I contient un élément inversible de A si et seulement si $I = A$. En particulier, si A est un corps, ses seuls idéaux sont $\{0\}$ et A.*

Démonstration. La condition est clairement nécessaire puisque, si $I = A$, $1 \in I$. Réciproquement, si $x \in I \cap A^*$, alors, en prenant $a = x^{-1}$ dans la définition d'un idéal, I contient $x^{-1}x = 1$ donc $y \cdot 1 = y$ pour tout $y \in A$. □

Exemple 66. *Les idéaux de $\mathbb{Z}$ sont les $n\mathbb{Z}$. En effet, on a déjà vu que les sous-groupes de $\mathbb{Z}$ sont les $n\mathbb{Z}$. Si $x \in n\mathbb{Z}$, c'est-à-dire $x = nq$ pour un $q \in \mathbb{Z}$, et si $a \in \mathbb{Z}$, alors $ax = naq \in n\mathbb{Z}$.*

Exemple 67. *Si A est un anneau, et si $x \in A$, l'ensemble*

$$I = Ax = \{ax : a \in A\}$$

est un idéal de A. En effet :

- *$0 \in I$ et, si $y = ax \in I$ et $y' = a'x \in I$, $y - y' = ax - a'x = (a - a')x \in I$. Donc $(I, +)$ est bien un sous-groupe de $(A, +)$.*
- *Si $y = ax \in I$ et si $b \in A$, $by = bax = (ba)x \in I$.*

On dit que I est un idéal principal.

Définition 68. *Soit A un anneau et I et J deux idéaux de A.*

(1) La somme $I + J$ de I et J est définie par :

$$I + J = \{x + y : x \in I, y \in J\}$$

(2) Le produit IJ de I et J est défini par :

$$IJ = \{\sum_{finie} xy : x \in I, y \in J\}$$

(3) L'intersection $I \cap J$ est l'intersection usuelle des ensembles.

Proposition 69. *Les ensembles $I + J$, IJ et $I \cap J$ sont des idéaux de A. En outre $IJ \subset I \cap J$, $I \cap J$ est le plus grand idéal de A contenu dans I et J, et $I + J$ est le plus petit idéal de A contenant I et J.*

Démonstration. Immédiat à partir de la définition d'un idéal. □

Définition 70. *Si S est une partie de A, on note (S) le plus petit idéal de A contenant S. On l'appelle* l'idéal engendré par S.

Proposition 71. *Si $S = \{x_1, \dots, x_s\}$,*

$$(S) = (x_1, \dots, x_s) = Ax_1 + \cdots + Ax_s.$$

III.4. Idéaux principaux, anneaux principaux. On commence par le définition suivante.

Définition 72. *Un* idéal principal *d'un anneau A est un idéal de la forme :*

$$I = Ax = \{ax : a \in A\}.$$

On dit que x est un générateur *de I. On note aussi $I = (x)$.*

Un anneau commutatif et intègre dont tous les idéaux sont principaux est appelé un anneau principal.

Exemple 73. *$\mathbb{Z}$ est un anneau principal.*

Exercice 74. *Montrez que, si $I = Ax$ est principal, les générateurs de I sont les ux avec $u \in A^*$.*

Les propriétés arithmétiques de $\mathbb{Z}$ s'étendent à un anneau A principal, et donc en particulier à $K[X]$, où K est un corps :

- Si $x, y \in A$, le pgcd de x et y est par définition un générateur de l'idéal $Ax + Ay$. Il est défini à la multiplication près par une unité de A. Dans le cas $A = K[X]$, on le choisit unitaire, c'est-à-dire de coefficient dominant 1, il est ainsi uniquement défini.
- On obtient par construction le théorème de Bezout : il existe u et v tels que $\text{pgcd}(x, y) = xu + yv$.
- Le ppcm de x et y est par définition un générateur de $Ax \cap Ay$. Dans le cas $A = K[X]$, on le choisit unitaire.
- On a $\text{pgcd}(x, y)\,\text{ppcm}(x, y) = xy$.
- La relation de divisibilité : $x \mid y$ s'il existe q tel que $y = qx$, est équivalente à la condition : $Ay \subset Ax$.
- La notion de nombre premier s'étend aussi : on parle d'irréductible d'un anneau pour un élément non inversible dont les seuls diviseurs sont 1 et lui-même, aux inversibles près. Alors, tout élément est le produit de façon essentiellement unique d'irréductibles de A.

Exercice 75. *Faire la liste des polynômes irréductibles de $\mathbb{Z}/2\mathbb{Z}[X]$ de degrés $1, 2, 3, 4$. Même question pour $\mathbb{Z}/3\mathbb{Z}[X]$.*

On a supposé dans tout ce chapitre que les anneaux considérés sont commutatifs. Dans le cas d'un anneau non commutatif, quelques nuances s'imposent : on distingue *idéaux à gauche* et *idéaux à droite*, vérifiant respectivement $a \in A, x \in I \Longrightarrow ax \in I$ et $a \in A, x \in I \Longrightarrow xa \in I$. Un idéal à droite et à gauche est appelé un *idéal bilatère*. Le quotient A/I est muni d'une structure d'anneau seulement si l'idéal est bilatère.

Exemple 76. *Dans $M_n(K)$, les seuls idéaux bilatères sont $\{0\}$ et $M_n(K)$. Par contre, il existe des idéaux non triviaux : par exemple, si V est un sous-espace vectoriel de K^n non trivial,*

$$I_V := \{M : xM = 0 \text{ pour tout } x \in V\},$$

est un idéal à droite.

Chapitre 2

Polynômes à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$

Ce chapitre est consacré à l'étude des polynômes à coefficients dans $K = \mathbb{R}$ ou $K = \mathbb{C}$.

I. Définitions et premières propriétés

Nous allons définir les polynômes à coefficients dans $K = \mathbb{R}$ ou $K = \mathbb{C}$ comme des suites à support fini (c'est-à-dire nulles à partir d'un certain rang) à valeurs dans K. Leur ensemble a une structure d'espace vectoriel, qui fait intervenir l'addition et la multiplication scalaire, et une opération supplémentaire de multiplication.

Définition 77. *Un* polynôme à coefficients dans K *est une suite finie $a_0, a_1, \ldots, a_n$ (mais de longueur arbitraire) d'éléments de K. On note un polynôme sous la forme*

$$P(X) = a_0 + a_1 X + \cdots + a_n X^n.$$

On dit que X est la variable *et que a_i est* le i-ème coefficient de P *ou* le coefficient de P de degré i. *Un polynôme de la forme $a_k X^k$ est appelé un* monôme.

On note $K[X]$ l'ensemble des polynômes à coefficient dans K.

On peut identifier $K[X]$ avec l'espace des suites à valeurs dans K et nulles à partir d'un certain rang (ou à support fini c'est pareil). Par conséquent, on connait déjà les opérations d'addition et de multiplication scalaire dans $K[X]$: pour additionner deux polynômes on additionne leurs coefficients de même degré et pour multiplier un polynôme par un scalaire $\lambda \in K$ on multiplie ses coefficients par λ.

On va maintenant introduire la multiplication de deux polynômes. On veut avoir $X^i.X^j = X^{i+j}$ et étendre aux polynômes par linéarité. Si $P = \sum a_i X^i$ et $Q = \sum b_i X^i$ alors le coefficient constant de PQ sera $a_0 b_0$, celui de degré 1 sera $a_0 b_1 + a_1 b_0$, celui de degré 2 sera $a_0 b_2 + a_1 b_1 + a_2 b_0$ et ainsi de suite. On est ainsi conduit à la définition suivante.

Définition 78. *Si $P = \sum_{k=0}^{n} a_k X^k$ et $Q = \sum_{k=0}^{m} b_k X^k$, alors on définit le polynôme PQ par : $PQ = \sum_{k=0}^{n+m} c_k X^k$ où*

$$c_k = \sum_{i+j=k} a_i b_j.$$

Proposition 79. *Les opérations d'addition et de multiplication sur $K[X]$ vérifient les propriétés suivantes :*

(1) $(K[X], +, \cdot)$ est un K-espace vectoriel

(2) Propriétés de la multiplication : pour tout $(P, Q, R) \in K[X]^3$,

- *(a) $0\,P = 0$ et $1\,P = P$*
- *(b) $PQ = QP$ (commutativité de la multiplication)*
- *(c) $(PQ)R = P(QR)$ (associativité de la multiplication)*
- *(d) $P(Q + R) = PQ + PR$ (distributivité de la multiplication sur l'addition)*

Démonstration. Laissée au lecteur. La démonstration des propriétés d'associativité et de distributivité est fastidieuse mais ne présente pas de difficultés. □

Exercice 80. *Démontrez* la formule du binôme de Newton *:*

$$(P+Q)^n = \sum_{k=0}^{n} \binom{n}{k} P^{n-k} Q^k$$

Définition 81. *Soit* $P = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$, $P \neq 0$. *Le* degré de P *est le plus grand indice* $k \geq 0$ *tel que* $a_k \neq 0$. *On le note* $\deg(P)$. *Par convention,* $\deg(0) = -\infty$. *Si* $k = \deg(P)$, *le* coefficient dominant *de* P *est le coefficient* a_k *et son* terme dominant *est* $a_k X^k$. *On dit que* P *est* unitaire *si son coefficient dominant vaut* 1. *On dit que* P *est* constant *s'il est de degré au plus* 0.

Proposition 82. *On a les propriétés suivantes (avec les conventions :* $-\infty + n = -\infty$, $-\infty - \infty = -\infty$, *etc..)*

(1) $\deg(PQ) = \deg(P) + \deg(Q)$

(2) $\deg(P+Q) \leq \deg(P) + \deg(Q)$

(3) Si $\deg(P) \neq \deg(Q)$ *alors* $\deg(P+Q) = \max(\deg(P), \deg(Q))$.

Remarque 83. *On note* $K_{n-1}[X]$ *l'ensemble des polynômes de degré strictement inférieur à* n. *C'est un* K*-espace vectoriel de base* $\{1, X, \ldots, X^{n-1}\}$ *et donc de dimension* n *sur* K.

La notion de degré d'un polynôme est très utile en arithmétique comme on le verra un peu plus tard avec la division Euclidienne. Pour l'instant, on va l'utiliser pour déterminer les polynômes inversibles.

Proposition-Définition 84. *On dit qu'un polynôme* $P \in K[X]$ *est* inversible dans $K[X]$ *s'il existe un polynôme* Q *tel que* $PQ = 1$. *Les seuls polynômes inversibles sont les polynômes constants non nuls.*

Démonstration. Supposons que $PQ = 1$. En prenant les degrés, on voit qu'on doit avoir $\deg(P) + \deg(Q) = 0$ ce qui impose $\deg(P) = \deg(Q) = 0$ donc P et Q sont constants non nuls. Dans ce cas, si $P = a \in K^*$, P a pour inverse $Q = a^{-1}$. □

On introduit maintenant une opération de dérivation formelle sur $K[X]$.

Définition 85. *Soit* $P(X) = \sum_{k=0}^{n} a_k X^k \in K[X]$, *on appelle* polynôme dérivé de P *et on note* P' *le polynôme défini par* $P'(X) = \sum_{k=1}^{n} a_k k X^{k-1}$.

Plus généralement, on note $P^{(k)}$ *le polynôme obtenu à partir de* P *après* k *dérivations. Par convention,* $P^{(0)} = P$, *et* $P^{(1)} = P'$.

La dérivation des polynômes vérifie les propriétés usuelles de la dérivation des fonctions :

Proposition 86. *On a les propriétés suivantes :*

(1) Si P *est constant,* $P' = 0$

(2) Si $\deg(P) > 0$, $\deg(P') = \deg(P) - 1$

(3) $(\lambda P + \mu Q)' = \lambda P' + \mu Q'$

(4) $(PQ)' = P'Q + PQ'$

(5) Plus généralement, on a la formule de Leibniz *:*

$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(n-k)} Q^{(k)}.$$

On laisse en exercice au lecteur la démonstration des formules usuelles de dérivation ; en particulier, la formule de dérivation de la 'composition' de deux polynômes $P(Q(X))$ (où on remplace dans P la variable X par le polynôme $Q(X)$).

$$\big(P(Q(X))\big)' = P'(Q(X))Q'(X).$$

Définition 87. *Soit* $P = \sum_{k=0}^{n} a_k X^k \in K[X]$ *et* $b \in K$. *L'*évaluation de P en b *est la valeur* $P(b) = \sum_{k=0}^{n} a_k b^k \in K$.

Proposition 88. *Soit $b \in K$, et P un polynôme de degré n. On a la* formule de Taylor *en b :*

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(b)}{k!}(X-b)^k.$$

Démonstration. Si $P(X) = a_0 + a_1 X + \cdots + a_n X^n$, on voit par dérivations successives que $a_0 = P(0)$, $a_1 = P'(0)$, et plus généralement que $a_k = P^{(k)}(0)/k!$. D'où la formule pour $b = 0$:

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k.$$

Pour obtenir la formule en b, on peut appliquer la formule en 0 au polynôme $Q(X) = P(X+b)$ en remarquant que $Q^{(k)}(X) = P^{(k)}(X+b)$, puis substituer $X - b$ à X. □

Définition 89. *On dit que $b \in K$ est* une racine de multiplicité m de P *si $P = (X-b)^m Q(X)$ avec $Q(b) \neq 0$.*

Théorème 90.

Soit $m \geq 1$ entier. On a équivalence de :

(1) b est une racine de P de multiplicité m,

(2) $P(b) = P'(b) = \cdots = P^{(m-1)}(b) = 0$ et $P^{(m)}(b) \neq 0$.

Démonstration. On commence par supposer que b est une racine de P de multiplicité m. Ainsi il existe Q dans $K[X]$ avec $Q(b) \neq 0$ tel que $P = (X-b)^m Q(X)$. On applique la formule de Leibniz avec $n \leqslant m$:

$$\begin{aligned} P^{(n)} &= \sum_{k=0}^{n} \binom{n}{k} ((X-b)^m)^{(k)} Q^{(n-k)} \\ &= \sum_{k=0}^{n} \binom{n}{k} m(m-1)\cdots(m-k+1)((X-b)^{m-k} Q^{(n-k)}, \end{aligned}$$

de sorte que pour $n < m$, l'indice k satisfait $k < m$ et donc $P^{(n)}(b) = 0$ alors que pour $n = m$, on obtient $P^{(m)}(b) = m!\, Q(b) \neq 0$.

Montrons maintenant la réciproque. On suppose que $P(b) = P'(b) = \cdots = P^{(m-1)}(b) = 0$ et $P^{(m)}(b) \neq 0$. On applique la formule de Taylor en b pour obtenir

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(b)}{k!}(X-b)^k,$$

et donc en particulier on obtient que $m \leqslant n$. Ainsi

$$\begin{aligned} P(X) &= \sum_{k=m}^{n} \frac{P^{(k)}(b)}{k!}(X-b)^k \\ P(X) &= (X-b)^m\, Q(X), \end{aligned}$$

where

$$Q(X) = \sum_{k=m}^{n} \frac{P^{(k)}(b)}{k!}(X-b)^{k-m}$$
$$= \sum_{j=0}^{n-m} \frac{P^{(m+j)}(b)}{(m+j)!}(X-b)^{j}.$$

En particulier $Q(b) = \frac{P^m(b)}{m!} \neq 0$. □

Définition 91. *Soit $P = \sum_{k=0}^{n} a_k X^k \in K[X]$ et $b \in K$. On dit que b* est une racine de P *si $P(b) = 0$.*

On observe que si b est une racine de multiplicité $m \geq 1$ de P alors b est une racine de P. On verra dans la prochaine section que la réciproque est vraie.

II. Division euclidienne (dans $\mathbb{Z}$ et) dans $K[X]$

On va maintenant aborder l'arithmétique sur les polynômes. On a sur les polynômes une division euclidienne tout à fait analogue à la division euclidienne des nombres entiers, et on va voir que les propriétés arithmétiques qui en découlent pour $K[X]$ sont essentiellement les mêmes que celles de $\mathbb{Z}$. Pour insister sur le parallèle des deux situations, on va les traiter dans ce qui suit simultanément. Dans le cas des entiers ce sont essentiellement des rappels de Lycée.

Commençons par la division euclidienne dans $\mathbb{Z}$.

Définition 92. *Pour tout couple d'entiers (a, b) avec $b \neq 0$, il existe un unique couple d'entiers (q, r) tels que*

$$a = bq + r \quad \textit{avec } 0 \leq r < |b|.$$

On dit que q est le quotient et que r est le reste de la division euclidienne de a par b.

De façon similaire on définit la division euclidienne dans $K[X]$.

Définition 93. *Pour tout couple de polynômes (A, B) avec $B \neq 0$, il existe un unique couple de polynômes (Q, R) tels que*

$$A = BQ + R \quad \textit{avec } \deg(R) < \deg(B).$$

On dit que Q est le quotient et que R est le reste de la division euclidienne de A par B.

Démonstration. On montre d'abord l'unicité : Si $A = BQ + R = BQ' + R'$ avec $\deg(R) < \deg(B)$ et $\deg(R') < \deg(B)$ alors $R' - R = B(Q - Q')$. Si $Q - Q' \neq 0$, on aurait

$$\deg(R' - R) = \deg(B) + \deg(Q - Q') \geq \deg(B),$$

ce qui contredirait les conditions $\deg(R), \deg(R') < \deg(B)$. Donc $Q = Q'$ et $R = R'$.

Montrons maintenant l'existence : tout d'abord, si $\deg(A) < \deg(B)$ on peut prendre $Q = 0$ et $R = A$. Supposons donc $n = \deg(A) \geq m = \deg(B)$. Alors $A' = A - B(a_n b_m^{-1} X^{n-m})$ est de degré strictement inférieur au degré de A. On procède alors par récurrence sur l'entier $k = n - m \geq 0$. Si $k = 0$, $Q = a_n b_m^{-1} X^{n-m}$ et $R = A'$ conviennent. Sinon, on applique l'hypothèse de récurrence à A' : il existe (Q', R') tels que $A' = BQ' + R'$ avec $\deg(R') < \deg(B)$. Alors, $A = B(a_n b_m^{-1} X^{n-m} + Q') + R'$ convient. □

Voyons une première conséquence sur les idéaux de $K[X]$. Rappelons qu'un idéal de $K[X]$ est un sous-ensemble non vide $\mathfrak{I}$ de $K[X]$ vérifiant les deux propriétés suivantes :

(a) $\forall Q_1 \in \mathfrak{I}, \forall Q_2 \in \mathfrak{I}, \quad Q_1 - Q_2 \in \mathfrak{I}$.

Proposition 94. *Soit $\mathfrak{I}$ un idéal de $K[X]$. Il existe un polynôme $\mathfrak{P}$ tel que $\mathfrak{I}$ soit l'ensemble des polynômes multiples de ce polynôme, i.e. $\mathfrak{I} = \{\mathfrak{P}Q,\ Q \in K[X]\}$. On dit que $\mathfrak{P}$* engendre $\mathfrak{I}$ *ou qu'il est un* générateur *de $\mathfrak{I}$.*

Démonstration. Si $\mathfrak{I} = \{0\}$ il suffit de prendre $\mathfrak{P} = 0$. Sinon, soit $\mathfrak{P}$ un polynôme non nul dans $\mathfrak{I}$ et de degré minimal. Comme $\mathfrak{I}$ est un idéal, il contient tous les multiples de $\mathfrak{P}$. Inversement si A est un élément quelconque de $\mathfrak{I}$, on fait la division euclidienne $A = \mathfrak{P}Q + R$. Comme $\mathfrak{I}$ est un idéal, $R = A - \mathfrak{P}Q$ appartient à $\mathfrak{I}$, et son degré est strictement inférieur à celui de $\mathfrak{P}$. Ainsi R est nul, à cause du choix de $\mathfrak{P}$, et A est donc un multiple de $\mathfrak{P}$. □

Remarque 95. *On peut remarquer qu'un idéal de $K[X]$, $\mathfrak{I} \neq \{0\}$, admet toujours un* unique générateur unitaire*, c'est-à-dire un générateur dont le coefficient dominant vaut 1. En effet si $\mathfrak{P}$ est un générateur de $\mathfrak{I}$ dont le coefficient dominant vaut $a \neq 0$ alors d'après la définition d'un idéal $\mathfrak{P}' = \mathfrak{P}/a$ appartient à $\mathfrak{I}$ et de plus est unitaire. Or, comme les degrés de $\mathfrak{P}'$ et $\mathfrak{P}$ sont égaux, $\mathfrak{P}'$ est encore un générateur de $\mathfrak{I}$. Maintenant si $\mathfrak{P}''$ est un autre générateur unitaire de $\mathfrak{I}$, alors $\mathfrak{P}''$ appartient à l'idéal $\mathfrak{I}$. Il existe donc $Q \in K[X]$ tel que $\mathfrak{P}'' = \mathfrak{P}'Q$. Mais comme $\mathfrak{P}''$ et $\mathfrak{P}'$ sont deux générateurs unitaires, ils ont même degré et même coefficient dominant. Donc $Q = 1$ et $\mathfrak{P}'' = \mathfrak{P}'$.*

Définition 96. *Soit $(a, b) \in \mathbb{Z}^2$.*

- *On dit que b* divise *a (ou que b* est un diviseur de *a ou que a* est un multiple de *b s'il existe $q \in \mathbb{Z}$ tel que $a = bq$. On note $b \mid a$.*
- *On dit que $d \in \mathbb{Z}$ est un* diviseur commun *de a et b si $d \mid a$ et $d \mid b$.*
- *Si $(a, b) \neq (0, 0)$, on dit que le pgcd (plus grand diviseur commun) de a et b est le plus grand entier d tel que $d \mid a$ et $d \mid b$. On le note* $\mathrm{pgcd}(a, b)$*. On pose* $\mathrm{pgcd}(0, 0) = 0$.

Définition 97. *Soit $(A, B) \in K[X]^2$.*

- *On dit que B* divise *A s'il existe $Q \in K[X]$ tel que $A = BQ$. On note $B \mid A$.*
- *On dit que $D \in K[X]$ est un* diviseur commun *de A et B si $D \mid A$ et $D \mid B$.*
- *Si $(A, B) \neq (0, 0)$, on dit qu'un pgcd de A et B est un polynôme D* de degré maximal *tel que $D \mid A$ et $D \mid B$. On note* $\mathrm{pgcd}(A, B)$ *l'ensemble des pgcd de A et B. On pose* $\mathrm{pgcd}(0, 0) = \{0\}$.

Remarque 98. *Quelques remarques sur ces notions dans $K[X]$:*

(1) Si B divise A, on peut écrire $A = BQ$ mais aussi, pour tout $\lambda \in K^*$, $A = (\lambda B)(\lambda^{-1}Q)$ de sorte que λB est aussi un diviseur de A. Pour cette raison, on considère le plus souvent seulement les diviseurs unitaires (de sorte qu'il y en ait un seul parmi les λB), de même que, dans $\mathbb{Z}$, on se restreint aux diviseurs positifs d'un entier.

(2) Remarquons que dans le cas des entiers, l'unicité du pgcd découle du fait que l'ensemble des diviseurs d'un entier non nul est fini (car borné) ; l'ensemble des diviseurs communs de a et b, où $(a, b) \neq (0, 0)$ possède donc un unique plus grand élément. Dans le cas des polynômes, ce raisonnement permet seulement de conclure que l'ensemble des degrés des diviseurs de A et B possède un plus grand élément ; il pourrait exister plusieurs diviseurs de même degré maximal et non proportionnels. On verra bientôt que ce n'est pas le cas mais il faudra travailler un peu pour obtenir cela.

Remarque 99. *Si $b = 0$, tout entier divise b, et pour tout $a \neq 0$,* $\mathrm{pgcd}(a, 0) = a$.
Si $B = 0$, tout polynôme divise B, et, pour tout $A \neq 0$, $\mathrm{pgcd}(A, B) = \{\lambda A \mid \lambda \in K\}$.

Définition 100. *Une* relation de Bezout *entre a et b deux entiers est une relation de la forme*

$$d = au + bv$$

où $d = \mathrm{pgcd}(a, b)$ et $(u, v) \in \mathbb{Z}^2$.

Définition 101. *Une* relation de Bezout *entre A et B deux polynômes est une relation de la forme*

$$D = AU + BV$$

où $D \in \mathrm{pgcd}(A, B)$ et $(U, V) \in K[X]^2$.

III. Algorithme d'Euclide étendu et relation de Bezout

Dans ce paragraphe on discute l'algorithme d'Euclide étendu, qui calcule efficacement un (le) pgcd et une relation de Bezout, de deux entiers ou de deux polynômes. En particulier cet algorithme montre de façon constructive *l'existence* d'une relation de Bezout. Soit donc $(A, B) \in \mathbb{Z}^2$ ou $(A, B) \in K[X]^2$. On suppose toujours que $(A, B) \neq (0, 0)$, et, sans perte de généralité, que $A \geq B$ (respectivement $\deg(A) \geq \deg(B)$).

Soit $R_0, R_1, \ldots, R_k, \ldots$ la suite des restes obtenus par divisions euclidiennes successives à partir de $R_0 = A$, $R_1 = B$, et ce tant que $R_k \neq 0$; on a donc pour tout $k \geq 1$

$$R_{k-1} = R_k Q_k + R_{k+1} \text{ avec } \begin{cases} 0 \leq R_{k+1} < R_k & \text{(cas de } \mathbb{Z}) \\ \deg(R_{k+1}) < \deg(R_k) & \text{(cas de } K[X]). \end{cases}$$

Remarquons d'abord que cette suite est finie, c'est-à-dire qu'au bout d'un nombre fini de divisions on obtient un reste nul. En effet, dans $\mathbb{Z}$ la suite R_k est une suite d'entiers positifs ou nuls strictement décroissante donc elle ne peut qu'atteindre 0 ; pour $K[X]$, on fait le même raisonnement avec la suite des degrés $\deg(R_k)$, qui va nécessairement atteindre $-\infty$.

Lemme 102. *Avec les notations précédentes, on a, pour tout $k \geq 1$, $\operatorname{pgcd}(R_{k-1}, R_k) = \operatorname{pgcd}(R_k, R_{k+1})$. En particulier, si n est le plus petit entier tel que $R_n \neq 0$, alors $R_n \in \operatorname{pgcd}(A, B)$.*

Démonstration. Soit D_k un pgcd de R_k et R_{k+1}, alors D_k divise R_k et R_{k+1}, donc il divise aussi $R_{k-1} = R_k Q_k + R_{k+1}$. Donc on peut conclure que, dans $\mathbb{Z}$, $D_k \leq D_{k-1}$, et dans $K[X]$, que $\deg(D_k) \leq \deg(D_{k-1})$. Mais on peut aussi faire le raisonnement inverse : $D_{k-1} \mid R_{k-1}$ et $D_{k-1} \mid R_k$ donc D_{k-1} divise $R_{k+1} = R_{k-1} - R_k Q_k$. Comme D_{k-1} divise R_k et R_{k+1}, on peut conclure que $D_{k-1} \leq D_k$ (respectivement que $\deg(D_{k-1}) \leq \deg(D_k)$). Dans le cas de $\mathbb{Z}$, on conclut immédiatement que $D_{k-1} = D_k$. Dans le cas de $K[X]$, on conclut que $\deg(D_{k-1}) = \deg(D_k)$ et donc que D_{k-1} est un pgcd de R_k et R_{k+1} et réciproquement. □

Revenons maintenant à la relation $R_{k-1} = R_k Q_k + R_{k+1}$. On peut exprimer cette relation par une relation matricielle :

$$\begin{pmatrix} R_k \\ R_{k+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -Q_k \end{pmatrix} \begin{pmatrix} R_{k-1} \\ R_k \end{pmatrix}.$$

En itérant cette formule, on obtient :

$$\begin{pmatrix} R_k \\ R_{k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & -Q_k \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -Q_{k-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -Q_1 \end{pmatrix}}_{M_k} \begin{pmatrix} R_0 \\ R_1 \end{pmatrix}$$

soit

$$\begin{pmatrix} R_k \\ R_{k+1} \end{pmatrix} = M_k \begin{pmatrix} A \\ B \end{pmatrix}.$$

On a pour les matrices M_k :

$$M_k = \begin{pmatrix} 0 & 1 \\ 1 & -Q_k \end{pmatrix} M_{k-1}$$

ce qui montre que l'on peut poser

$$M_k = \begin{pmatrix} U_k & V_k \\ U_{k+1} & V_{k+1} \end{pmatrix}$$

avec

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{cases} U_{k+1} = U_{k-1} - U_k Q_k \\ V_{k+1} = V_{k-1} - V_k Q_k \end{cases}.$$

Finalement, on a pour $k = n - 1$,

$$\begin{pmatrix} R_{n-1} \\ R_n \end{pmatrix} = \begin{pmatrix} U_{n-1} & V_{n-1} \\ U_n & V_n \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}.$$

et donc

$$R_n = AU_n + BV_n$$

qui est une relation de Bezout entre A et B.

On résume nos conclusions sous la forme d'un algorithme calculant un pgcd et une relation de Bezout de A et B dans le théorème suivant :

Théorème 103.

Soit $(A, B) \in \mathbb{Z}^2$ tels que $A \geq B$ ou $(A, B) \in K[X]^2$ tels que $\deg(A) \geq \deg(B)$. On suppose $(A, B) \neq (0, 0)$. Soit R_k, Q_k, U_k et V_k les entiers ou polynômes calculés itérativement à partir des initialisations :

$$R_0 = A \quad U_0 = 1 \quad V_0 = 0$$
$$R_1 = B \quad U_1 = 0 \quad V_1 = 1$$

par division euclidienne :

$$R_{k-1} = R_k Q_k + R_{k+1}$$

où $0 \leq R_{k+1} < R_k$ (cas de $\mathbb{Z}$) ou $\deg(R_{k+1}) < \deg(R_k)$ (cas de $K[X]$), et les formules

$$U_{k+1} = U_{k-1} - U_k Q_k$$
$$V_{k+1} = V_{k-1} - V_k Q_k$$

jusqu'à l'obtention d'un reste nul. Soit R_n le dernier reste non nul. Alors, R_n est un pgcd de A et B et $R_n = AU_n + BV_n$ est une relation de Bezout.

Exemple 104. *Un exemple d'exécution de l'algorithme d'Euclide étendu sur les polynômes $A = X^4 - 1$, $B = X^3 - 2X + 1$.*

R_k	U_k	V_k	Q_k	
$X^4 - 1$	1	0		
$X^3 - 2X + 1$	0	1	X	$A = BX + (2X^2 - X - 1)$
$2X^2 - X - 1$	1	$-X$	$(2X + 1)/4$	$B = R_2(X/2 + 1/4) - 5/4(X - 1)$
$-5/4(X - 1)$	$-(2X + 1)/4$	$(2X^2 + X + 4)/4$	$-4/5(2X + 1)$	$R_2 = -4/5R_3(2X + 1)$
0				

On obtient que $(X - 1)$ est un pgcd de A et B, et la relation de Bezout :

$$-5(X - 1) = -(2X + 1)A + (2X^2 + X + 4)B.$$

Corollaire 105. *Le pgcd de deux polynômes est unique à multiplication par un scalaire près.*

Démonstration. Soit $(A, B) \neq (0, 0)$ deux polynômes. L'algorithme d'Euclide étendu calcule un pgcd D de A et B et deux polynômes U et V tels que $D = AU + BV$. Soit D' un autre pgcd de A et B. Alors D' divise A et B donc D' divise $AU + BV$ donc D' divise D. Il existe donc Q tel que $D = D'Q$. Mais, puisque D et D' sont tous les deux des pgcd, ils sont de même degré maximal, donc Q est de degré 0 soit une constante. □

Remarque 106. *Désormais, on définit LE pgcd de deux polynômes A et B, $(A, B) \neq (0, 0)$ comme étant l'unique polynôme* unitaire *appartenant à* $\text{pgcd}(A, B)$.

IV. Conséquences du théorème de Bezout

On a démontré, de façon constructive, au paragraphe précédent le théorème de Bezout :

Théorème 107.

Soit $(A, B) \in \mathbb{Z}^2$ (respectivement $(A, B) \in K[X]^2$), et soit $D = \text{pgcd}(A, B)$. Il existe $(U, V) \in \mathbb{Z}^2$ (respectivement $(U, V) \in K[X]^2$) et D un pgcd de A et B tels que

$$D = AU + BV.$$

Démonstration. Si $A = B = 0$ on peut prendre $U = V = 0$ puisqu'on a posé $\text{pgcd}(A, B) = 0$. Si $(A, B) \neq (0, 0)$, l'algorithme d'Euclide étendu calcule un multiple scalaire du pgcd de A et B et une relation de Bezout pour ce multiple. □

On va maintenant en déduire quelques conséquences importantes.

Proposition 108. *Soit $(A, B) \in \mathbb{Z}^2$ ou $(A, B) \in K[X]^2$. Soit $D = \text{pgcd}(A, B)$. Alors on a*

$$C \mid A \text{ et } C \mid B \iff C \mid D.$$

Démonstration. Seule l'implication $\Rightarrow$ n'est pas triviale. Pour la démontrer, on utilise une relation de Bezout $D = AU + BV$. Si $C \mid A$ et $C \mid B$, alors $C \mid (AU + BV)$ donc $C \mid D$. □

Définition 109. *Soit $(A, B) \in \mathbb{Z}^2$ ou $(A, B) \in K[X]^2$. On dit que A et B sont* premiers entre eux *si* $\text{pgcd}(A, B) = 1$.

Proposition 110. *Soit $(A, B) \in \mathbb{Z}^2$ ou $(A, B) \in K[X]^2$.*

(1) A et B sont premiers entre eux si et seulement s'il existe (U, V) tels que $1 = AU + BV$.

(2) Si $D = \text{pgcd}(A, B)$ alors il existe $(A', B') \in \mathbb{Z}^2$ (respectivement $(A', B') \in K[X]^2$ tels que $A = DA'$, $B = DB'$ et A' et B' sont premiers entre eux.

Démonstration. Démontrons 1. : si $\text{pgcd}(A, B) = 1$, le théorème de Bezout montre l'existence de (U, V) tels que $1 = AU + BV$ (noter que la relation de Bezout obtenue par l'algorithme d'Euclide étendu doit éventuellement être multipliée par une constante non nulle). Réciproquement, si on a $1 = AU + BV$, $D = \text{pgcd}(A, B)$ divise A et B donc divise $AU + BV = 1$. Donc $D = 1$.

Démontrons 2. : Si $(A, B) = (0, 0)$, on peut prendre $A' = B' = 1$; on suppose maintenant $(A, B) \neq (0, 0)$. Si $D = \text{pgcd}(A, B)$ alors il existe $(A', B') \in \mathbb{Z}^2$ (respectivement $(A', B') \in K[X]^2$) tels que $A = DA'$, $B = DB'$. On remplace dans une relation de Bezout $D = AU + BV = D(A'U + B'V)$ d'où on tire $(D \neq 0)$ $1 = A'U + B'V$; par le résultat précédent, A' et B' sont donc premiers entre eux. □

Exemple 111. *Soit $(a, b) \in K^2$ avec $a \neq b$, et soit $A = X - a$, $B = X - b$. Il est clair que A et B n'ont pas de diviseurs communs autres que les polynômes constants. Donc* $\text{pgcd}(A, B) = 1$. *On a la relation de Bezout :* $1 = (B - A)/(b - a)$.

Le *Lemme de Gauss* est une conséquence importante du théorème de Bezout :

Théorème 112.

[Lemme de Gauss] Soit $(A, B, C) \in \mathbb{Z}^3$ ou $(A, B, C) \in K[X]^3$. Si $A \mid BC$ et si $\text{pgcd}(A, B) = 1$ *alors $A \mid C$.*

Démonstration. Soit $1 = AU + BV$ une relation de Bezout entre A et B. On multiplie par C pour obtenir $C = ACU + BCV$. Comme $A \mid ACU$ et $A \mid BCV$ alors $A \mid (ACU + BCV)$ et donc $A \mid C$. □

Corollaire 113. *Si $A \mid C$, si $B \mid C$, et si* $\operatorname{pgcd}(A,B) = 1$*, alors $AB \mid C$.*

Démonstration. Puisque $A \mid C$, il existe Q tel que $C = AQ$. On a $B \mid C = AQ$ et $\operatorname{pgcd}(A,B) = 1$ donc par le Lemme de Gauss, $B \mid Q$. Il existe donc Q' tel que $Q = BQ'$; alors $C = AQ = ABQ'$ ce qui montre que $AB \mid C$. □

V. Polynômes irréductibles

Dans ce paragraphe, on va démontrer l'analogue pour les polynômes de la factorisation des entiers en produit de nombres premiers. Le rôle des nombres premiers va être joué par *les polynômes irréductibles.*

Définition 114. *Soit $P \in K[X]$. On dit que P est* un polynôme irréductible *si* $\deg(P) \geq 1$ *et si on ne peut pas le factoriser sous la forme d'un produit de deux polynômes de $K[X]$ de degrés plus grand que* 0*.*

Autrement dit, un polynôme P est irréductible si P n'est pas constant et si $\operatorname{pgcd}(P,A) = 1$ pour tout $A \neq 0$. C'est clairement le cas de tous les polynômes de degré 1. En effet, si $\deg(P) = 1$ et si $P = AB$ alors, en prenant les degrés, $1 = \deg(A) + \deg(B)$ ce qui impose $\deg(A) = 0$ ou $\deg(B) = 0$, soit A ou B est constant.

Si $P = aX^2 + bX + c$, $a \neq 0$, et si P a une factorisation non triviale (c'est-à-dire par des polynômes non constants), ses facteurs doivent être de degré 1. Donc P se factorise si et seulement si P possède deux racines (éventuellement égales). On sait que, si $K = \mathbb{C}$, c'est toujours le cas, et si $K = \mathbb{R}$, c'est le cas si et seulement si son discriminant $\Delta = b^2 - 4ac$ est positif ou nul.

Remarque 115. *La propriété d'être irréductible dépend vraiment du corps K que l'on considère. Ainsi, $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$ mais pas dans $\mathbb{C}[X]$ puisque $X^2 + 1 = (X - i)(X + i)$.*

On a donc caractérisé les polynômes irréductibles de $K[X]$ de degré 1 et 2 ; en fait il n'y en a pas d'autres mais pour prouver cela on doit s'appuyer sur un théorème difficile et profond, qu'on admettra dans le cadre de ce cours, c'est le *théorème de d'Alembert-Gauss* :

Théorème 116.

[Théorème de d'Alembert-Gauss] Tout polynôme de $\mathbb{C}[X]$ de degré au moins égal à 1 admet au moins une racine.

Corollaire 117. *Soit $P \in \mathbb{C}[X]$, $\deg(P) = n \geq 1$, de coefficient dominant a_n. Il existe $(z_1, \ldots, z_k) \in \mathbb{C}^k$, $z_i \neq z_j$, et $(m_1, \ldots, m_k) \in \mathbb{N}^k$, $m_i > 0$, tels que*

$$P(X) = a_n (X - z_1)^{m_1} \ldots (X - z_k)^{m_k}.$$

Les zéros de P sont les z_i, les m_i sont leurs multiplicités, et $\sum_{i=1}^{k} m_k = n$ (autrement dit : le nombre de racines de P comptées avec leurs multiplicités est égal au degré de P).

Démonstration. Par récurrence sur $n = \deg(P)$. Le théorème de d'Alembert-Gauss montre que P possède au moins une racine, notons-la z_1 ; par division Euclidienne, $P = (X - z_1)Q$, et on peut appliquer l'hypothèse de récurrence à Q. □

Corollaire 118. *Soit $P \in \mathbb{R}[X]$, $\deg(P) = n \geq 1$, de coefficient dominant a_n. Il existe $P_1, \ldots, P_k$, $P_i \neq P_j$, avec $P_i = (X - x_i)$, $x_i \in \mathbb{R}$ ou $P_i = X^2 + b_i X + c_i$, $b_i^2 - 4c_i < 0$, et $(m_1, \ldots, m_k) \in \mathbb{N}^k$, $m_i > 0$, tels que*

$$P(X) = a_n P_1^{m_1} \ldots P_k^{m_k}.$$

Les x_i sont les racines réelles de P, de multiplicité m_i, et les racines complexes des polynômes P_i de degré 2 sont les racines complexes conjuguées de P, elles sont de multiplicité m_i.

Démonstration. Comme $P \subset \mathbb{R}[X]$ et $\mathbb{R} \subset \mathbb{C}$, on peut factoriser P dans $\mathbb{C}[X]$ en appliquant le résultat précédent. Il existe donc $(z_1, \ldots, z_k) \in \mathbb{C}^k$, $z_i \neq z_j$, et $(m_1, \ldots, m_k) \in \mathbb{N}^k$, $m_i > 0$, tels que

$$P(X) = a_n(X - z_1)^{m_1} \ldots (X - z_k)^{m_k}.$$

On remarque ensuite que, si z est une racine de P qui n'est pas réelle, alors, comme $P \in \mathbb{R}[X]$, son conjugué $\overline{z}$ est aussi racine de P, qui plus est avec la même multiplicité. En effet, si $P = (X - z)^m Q$ avec $Q(z) \neq 0$ alors en conjuguant les coefficients, on obtient $P = (X - \overline{z})^m \overline{Q}$ et $\overline{Q}(\overline{z}) \neq 0$. On peut donc réorganiser les racines de P en racines réelles $x_1, \ldots, x_\ell$, et paires de racines complexes conjuguées $(z_{\ell+1}, \overline{z_{\ell+1}}), \ldots, (z_k, \overline{z_k})$, puis poser

$$\begin{cases} P_i = (X - x_i) & 1 \leq i \leq \ell \\ P_i = (X - z_i)(X - \overline{z_i}) & \ell + 1 \leq i \leq k. \end{cases}$$

Les polynômes P_i sont tous à coefficients réels, en effet, $(X - z)(X - \overline{z}) = X^2 - (z + \overline{z})X + z\overline{z} \in \mathbb{R}[X]$ et, si $z \notin \mathbb{R}$, ce polynôme a un discriminant négatif, donc on a comme annoncé

$$P = P_1^{m_1} \ldots P_k^{m_k}$$

avec les conditions requises sur les polynômes P_i. □

Chapitre 3

Réduction des endomorphismes

Pour décrire un endomorphisme f d'un espace vectoriel E, on cherche une base de E dans laquelle la matrice de f soit la plus simple possible. Pour diverses raisons, on voudrait que cette matrice soit *diagonale*, c'est-à-dire que les coefficients en dehors de la diagonale soient nuls. En d'autre termes, les vecteurs de cette base sont tels que leur image par f leur est colinéaire. On les appelle des *vecteurs propres* de f. L'avantage d'avoir une matrice diagonale est qu'il est alors plus facile de faire des calculs faisant intervenir f, par exemple le calcul des itérés de f.

Il n'existe pas toujours de base de E formée de vecteurs propres de f. Par exemple la rotation R_α dans le plan réel orienté, de matrice

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

dans la base canonique, n'a pas de vecteur propre si $\alpha \not\equiv 0 \mod \pi$, puisqu'aucun vecteur non nul x n'est colinéaire à $R_\alpha(x)$. Le problème dans ce cas se résout en passant du corps $\mathbb{R}$ au corps $\mathbb{C}$: nous verrons qu'un endomorphisme d'un espace vectoriel complexe a toujours un vecteur propre. Néanmoins, même sur $\mathbb{C}$, il peut ne pas y avoir "assez" de vecteurs propres pour former une base ; c'est par exemple le cas pour l'endomorphisme de $\mathbb{C}^2$ de matrice

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

dont tous les vecteurs propres sont colinéaires à e_1 et donc ne peuvent pas former une base. On cherchera pour ces endomorphismes une base dans laquelle leur matrice est aussi simple que possible, c'est-à-dire aussi proche que possible d'une forme diagonale.

Dans tout ce chapitre, les espaces vectoriels considérés sont des espaces vectoriels sur K de dimension finie, où K est le corps des nombres réels $\mathbb{R}$ ou le corps des nombres complexes $\mathbb{C}$. En fait les résultats sont encore vrais si K est un corps commutatif contenant $\mathbb{Q}$.

I. Rappel sur les espaces vectoriels

Il sera utile pour la suite de ce chapitre d'être au point sur la notion d'espace vectoriel et en particulier de somme directe d'espaces vectoriels.

I.1. Définitions. Nous commençons par rappeler la notion d'espace vectoriel.

Définition 119. *Soit E un ensemble non vide, muni de deux opérations : une opération (ou loi) interne appelée* addition *et notée $+$:*

$$E \times E \to E \qquad (u, v) \mapsto u + v$$

et une opération (ou loi) externe appelée multiplication *et notée $\cdot$ de $K \times E$ vers E qui à (λ, u) associe $\lambda \cdot u$.*

On dit que $(E, +, \cdot)$ est un espace vectoriel sur K *ou un* K-espace vectoriel *si les propriétés suivantes sont vérifiées :*

- *L'addition sur E est :*
  - *(1) commutative : $u + v = v + u$ pour tout $(u, v) \in E^2$*
  - *(2) associative : $(u + v) + w = u + (v + w)$ pour tout $(u, v, w) \in E^3$*
  - *(3) possède un zéro : il existe $\mathbf{0} \in E$ tel que $u + \mathbf{0} = \mathbf{0} + u = u$, pour tout $u \in E$*
  - *(4) tout élément possède un opposé : pour tout $u \in E$, il existe un élément noté $-u$ tel que $u + (-u) = \mathbf{0}$.*

- *La multiplication externe vérifie :*

  (5) $1 \cdot u = u$ *pour tout* $u \in E$

  (6) $\lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u$ *pour tout* $(\lambda, \mu) \in K^2$, $u \in E$.

- *La multiplication est* distributive sur l'addition *:*

  (7) $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$ *pour tout* $\lambda \in K$, *pour tout* $(u, v) \in E^2$

  (8) $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot v$ *pour tout* $(\lambda, \mu) \in K^2$, *pour tout* $u \in E$.

On appelle les éléments d'un espace vectoriel des *vecteurs* et les éléments de K des *scalaires*. L'élément $\mathbf{0}$ est appelé le *vecteur nul*.

Les opérations d'espaces vectoriels vérifient quelques propriétés supplémentaires qui se déduisent de celles listées dans la définition :

(1) $\lambda \cdot \mathbf{0} = \mathbf{0}$ pour tout $\lambda \in K$,

(2) $0 \cdot u = \mathbf{0}$ pour tout $u \in E$,

(3) Si $\lambda \cdot u = \mathbf{0}$ alors $\lambda = 0$ ou $u = \mathbf{0}$,

(4) $-u = (-1) \cdot u$ pour tout $u \in E$,

(5) Si $\lambda_1, \ldots, \lambda_k$ appartiennent à K, et si $v_1, v_2, \ldots, v_k$ appartiennent à E, alors

$$u = \lambda_1 v_1 + \cdots + \lambda_n v_k$$

appartient à E. On dit que u est une *combinaison linéaire* des vecteurs $v_1, \ldots, v_k$.

Par exemple K^n est un K-espace vectoriel, ainsi que l'espace vectoriel nul $\{\mathbf{0}\}$ où $\mathbf{0}$ est un 'symbole'. Une manière de multiplier les exemples est d'appeler la définition suivante.

Définition 120. *Soit* $(E, +, \cdot)$ *un espace vectoriel sur* K. *Un sous-espace vectoriel de* E *est un sous-ensemble* F *de* E *vérifiant :*

(1) $F \neq \emptyset$

(2) Pour tout $(u, v) \in F^2$, $u + v \in F$

(3) Pour tout $\lambda \in K$, $u \in F$, $\lambda u \in F$.

En effet il est facile de vérifier que si F est un sous-espace vectoriel de $(E, +, \cdot)$ alors $(F, +, \cdot)$ est lui-même un espace vectoriel sur K. Il est d'ailleurs souvent préférable, pour montrer qu'un certain ensemble est un espace vectoriel, de montrer que c'est un sous-espace vectoriel d'un espace vectoriel déjà connu que de vérifier tous les axiomes.

I.2. Somme et somme directe de sous-espaces vectoriels. Soit $(E, +, \cdot)$ un espace vectoriel et soit U et V deux sous-espaces vectoriels de E. On rappelle la notion de somme $U + V$: c'est le sous-espace vectoriel de E défini par :

$$U + V = \{u + v \mid u \in U, v \in V\}.$$

L'intersection $U \cap V$ est aussi un sous-espace vectoriel de E et on a, si E est de dimension finie, la formule dite de Grassmann :

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V).$$

Définition 121. *On dit que* E *est* la somme directe de U et V, *et on note* $E = U \oplus V$, *si les conditions suivantes sont réalisées :*

(1) $E = U + V$

(2) $U \cap V = \{\mathbf{0}\}$

Théorème 122.

Les conditions suivantes sont équivalentes :

(1) $E = U \oplus V$

(2) Tout vecteur $u \in E$ *s'écrit* d'une façon unique $u = x + y$ *avec* $x \in U$ *et* $y \in V$.

Si, en outre, E *est de dimension finie, alors on a équivalence de :*

(3) $E = U \oplus V$

(4) $E = U + V$ *et* $\dim(E) = \dim(U) + \dim(V)$

(5) $U \cap V = \{0\}$ *et* $\dim(E) = \dim(U) + \dim(V)$.

Démonstration. Montrons d'abord l'équivalence de (1) et (2). Supposons $E = U \oplus V$, et soit $x \in E$. Comme $E = U + V$, on sait qu'il existe $u \in U$, $v \in V$ tels que $x = u + v$. S'il y avait deux décompositions de x, on aurait $x = u + v = u' + v''$ mais alors $u - u' = v' - v \in U \cap V$. L'hypothèse $U \cap V = \{\mathbf{0}\}$ montre que $u - u' = v' - v = \mathbf{0}$ donc $u = u'$ et $v = v'$.

Réciproquement, si (2) est vrai, alors on a bien sûr $E = U + V$. Il reste à montrer que $U \cap V = \{0\}$. Si $x \neq \mathbf{0}$, $x \in U \cap V$, on aurait deux décompositions de $\mathbf{0}$ en la somme d'un vecteur de U et d'un vecteur de V : $\mathbf{0} = \mathbf{0} + \mathbf{0} = x + (-x)$ ce qui contredirait (2).

Les équivalences de (3), (4), (5), se montrent avec la formule $\dim(U + V) = \dim(U) + \dim(V) = \dim(U \cap V)$. □

Exercice : Montrez que, si $E = U \oplus F$, la réunion d'une base de U et d'une base de V est une base de E.

II. Déterminant

II.1. Définition et premières propriétés du déterminant d'une matrice. Dans cette section nous allons définir successivement le déterminant d'une matrice, d'une famille de vecteurs et enfin d'un endomorphisme. Nous établirons, chemin faisant, les propriétés essentielles du déterminant. Nous verrons notamment que le déterminant permet, dans certains cas, de montrer si une application linéaire est inversible ou non. Il permettra aussi, toujours dans certains cas, de résoudre des systèmes ou bien d'obtenir l'inverse d'une matrice. Enfin il servira à la diagonalisation et la trigonalisation des endomorphismes d'un espace vectoriel.

On commence par définir le déterminant d'une matrice.

Définition 123. *Soit* $M = (m_{ij})_{1 \leq i,j \leq n}$ *une matrice carrée d'ordre* n *à coefficients dans un corps* K. *On appelle* déterminant de la matrice M *et on note* $\det(M)$ *le scalaire*

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(j)}$$

Remarque 124. *On note en général le déterminant de la matrice* M *:*

$$\det(M) = \begin{vmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ldots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{vmatrix}.$$

Exemple 125. *On a* $S_2 = \{\mathrm{Id},\ (1\ 2)\}$ *et donc :*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Exemple 126. *On a* $S_3 = \{\mathrm{Id},\ (1\ 2),\ (1\ 3),\ (2\ 3),\ (1\ 2)(1\ 3),\ (1\ 2)(2\ 3)\}$ *et donc*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33}.$$

Proposition 127.
a) *Si* $M = (m)$ *est une matrice d'ordre* 1 *alors* $\det(M) = m$.
b) *Si on note* I_n *la matrice identité d'ordre* n *alors* $\det(\mathrm{I}_n) = 1$.
c) *Pour tout* $\lambda \in K$ *et toute matrice carrée* M *d'ordre* n, *on a* $\det(\lambda M) = \lambda^n \det(M)$.

Démonstration. Le premier point est évident.
Pour le second, on remarque que si on note a_{ij} les coefficients de la matrice identité, alors $a_{j\sigma(j)} = 0$ si $\sigma(j) \neq j$, et donc $\det(\mathrm{I}_n) = \prod_{j=1}^n a_{jj} = 1$.
Enfin pour le troisième point si $M = (m_{ij})$ alors $\lambda M = (\lambda m_{ij})$ et donc

$$\det(\lambda M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n \lambda m_{j\sigma(j)} = \lambda^n \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n m_{j\sigma(j)} = \lambda^n \det(A).$$

□

Proposition 128. *On a aussi*

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n m_{\sigma(j)j}.$$

Démonstration. La preuve repose essentiellement sur deux changements de variable : un premier dans le produit et un second dans la somme. Soit ρ et σ dans S_n, alors comme ρ est une bijection de $\{1, \ldots, n\}$, en posant $j = \rho(k)$ on a

$$\prod_{j=1}^n m_{j\sigma(j)} = \prod_{k=1}^n m_{\rho(k)\sigma(\rho(k))}.$$

En particulier avec $\rho = \sigma^{-1}$, on a $\sigma(\rho(k)) = k$ et donc

$$\prod_{j=1}^n m_{j\sigma(j)} = \prod_{k=1}^n m_{\sigma^{-1}(k)k}.$$

En revenant à la définition du déterminant on obtient, par sommation, que

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{k=1}^n m_{\sigma^{-1}(k)k}.$$

On se rappelle ensuite que $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ de sorte que

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) \prod_{k=1}^n m_{\sigma^{-1}(k)k}.$$

Comme $\sigma \to \sigma^{-1}$ est une bijection de S_n, en posant $\rho = \sigma^{-1}$ on obtient

$$\det(M) = \sum_{\rho \in S_n} \varepsilon(\rho) \prod_{k=1}^n m_{\rho(k)k},$$

et on renomme la variable de sommation σ, au lieu de ρ, pour conclure. □

Corollaire 129. *Une matrice carrée et sa transposée ont des déterminants égaux.*

Démonstration. Soit $M = (m_{ij})$ une matrice carrée. On note ${}^tM = (m^*_{ij}) = (m_{ji})$ sa transposée. On a :

$$\det({}^tM) = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m^*_{j\sigma(j)} = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j} = \det(M),$$

d'après la proposition précédente. □

Remarque 130. *Cela signifie que toutes les propriétés du déterminant qui seront établies par rapport aux colonnes d'une matrice seront aussi valables pour les lignes.*

II.2. Determinant des matrices triangulaires par blocs. Nous allons voir que pour certaines matrices le calcul du déterminant se fait assez facilement.
Considérons tout d'abord des matrices dont l'écriture par blocs est

$$M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix},$$

où M' est une matrice carrée d'ordre p, M'' est une matrice carrée d'ordre q, N est une matrice à p lignes et q colonnes et le dernier bloc (noté 0 dans la matrice) est la matrice nulle à q lignes et p colonnes, avec $p + q = n$ l'ordre de M.

Proposition 131. *Le déterminant de la matrice* $M = \begin{pmatrix} M' & N \\ 0 & M'' \end{pmatrix}$ *est* $\det(M')\det(M'')$.

Démonstration. Notons $M = (m_{ij})_{1\le i,j\le n}$. On a

$$\det(M) = \sum_{\sigma\in S_n} \varepsilon(\sigma) m_{\sigma(1)1} \dots m_{\sigma(n)n}.$$

Pour qu'un terme soit non nul il faut que

$$\sigma(1), \dots, \sigma(p) \in \{1, \dots, p\},$$

et donc

$$\sigma(p+1), \dots, \sigma(n) \in \{p+1, \dots, n\}.$$

Il existe donc des permutations $\sigma' \in S_p$ et $\sigma'' \in S_q$ telle que

$$\sigma(j) = \sigma'(j) \quad \text{si} \quad 1 \le j \le p,$$
$$\sigma(p+k) = p + \sigma''(k) \quad \text{si} \quad 1 \le k \le q.$$

On a donc

$$\det(M) = \sum_{\sigma'\in S_p, \sigma''\in S_q} \varepsilon(\sigma) m'_{\sigma'(1)_1} \dots m'_{\sigma'(p)p} m''_{\sigma''(p+1)(p+1)} \dots m''_{\sigma''(n)n}.$$

En décomposant σ' et σ'' en produit de transpositions on remarque que $\varepsilon(\sigma) = \varepsilon(\sigma')\varepsilon(\sigma'')$, et donc

$$\det(M) = \left(\sum_{\sigma'\in S_p} \varepsilon(\sigma') m'_{\sigma'(1)_1} \dots m'_{\sigma'(p)p} \right) \left(\sum_{\sigma''\in S_q} \varepsilon(\sigma'') m''_{\sigma''(1)_1} \dots m''_{\sigma''(p)p} \right) = \det(M')\det(M'')$$

ce qui montre la proposition. □

Exemple 132. *En combinant les propriétés que l'on a vues, on obtient assez facilement que*

$$\begin{vmatrix} 11 & 0 & 12 & 0 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 7 & 5 & 8 & 6 \end{vmatrix} = - \begin{vmatrix} 7 & 5 & 8 & 6 \\ 3 & 1 & 4 & 2 \\ 9 & 0 & 10 & 0 \\ 11 & 0 & 12 & 0 \end{vmatrix} = \begin{vmatrix} 6 & 5 & 8 & 7 \\ 2 & 1 & 4 & 3 \\ 0 & 0 & 10 & 9 \\ 0 & 0 & 12 & 11 \end{vmatrix}$$
$$= \begin{vmatrix} 6 & 5 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} 10 & 9 \\ 12 & 11 \end{vmatrix} = (6.1 - 5.2)(10.11 - 9.12) = -8.$$

Une simple généralisation de ce résultat donne :

Corollaire 133. *Soit M une matrice de la forme*

$$\begin{pmatrix} M_1 & * & * & * \\ 0 & M_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & M_r \end{pmatrix}$$

où, pour i compris entre 1 et r, M_i est une matrice carrée d'ordre k_i, avec $k_1 + \cdots + k_r = n$. Alors on a

$$\det(M) = \det(M_1) \dots \det(M_r)$$

On en déduit alors le déterminant d'une matrice triangulaire supérieure. Rappelons qu'une matrice carrée d'ordre n est dite *triangulaire supérieure (inférieure)* si tous ses coefficients en dessous (respectivement au-dessus) de sa diagonale sont nuls.

Corollaire 134. *Soit M une matrice triangulaire supérieure, c'est-à-dire*

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m_{nn} \end{pmatrix},$$

alors $\det(M) = m_{11} \dots m_{nn}$

Remarque 135. *Le résultat est identique pour une matrice diagonale (matrice triangulaire supérieure dont les coefficients au-dessus de la diagonale sont tous nuls). D'autre part il reste encore vrai pour les matrices triangulaires inférieures (transposées de matrices triangulaires supérieures).*

II.3. Déterminant d'une famille de vecteurs, multiplicativite et critère d'inversibilité. On définit maintenant le déterminant d'une famille de vecteurs.

Définition 136. *Soient $(x_1, \dots, x_n)$ des vecteurs d'un K espace vectoriel E de dimension n. On appelle* déterminant *de $x_1, \dots, x_n$ dans une base $\mathcal{B}$ et on note $\det_{\mathcal{B}}(x_1, \dots, x_n)$ (on omet l'indice $\mathcal{B}$ lorsqu'il n'y a pas ambiguité) le déterminant de la matrice dont les colonnes sont constituées des vecteurs $x_1, \dots, x_n$.*

Proposition 137. *Soient $(x_1, \dots, x_n)$ des vecteurs d'un K espace vectoriel E de dimension n.*
a) *Soit $\rho \in S_n$, alors*

$$\det(x_{\rho(1)}, \dots, x_{\rho(n)}) = \varepsilon(\rho)\det(x_1, \dots, x_n).$$

En particulier si on échange deux colonnes d'une matrice, le déterminant est changé en son opposé.
b) *Si $x_i = x_j$ pour $i \neq j$ alors on a* $\det(x_1, \dots, x_n) = 0$.
c) *Soient y_k un vecteur de E et λ et μ deux scalaires, alors :*

$$\det(x_1, \dots, \lambda x_k + \mu y_k, \dots, x_n) = \lambda\det(x_1 \dots, x_k, \dots, x_n) + \mu\det(x_1 \dots, y_k, \dots, x_n).$$

Le déterminant d'une matrice dépend donc linéairement de chacun de ses vecteurs colonnes. En particulier il est nul si une colonne est nulle.
d) *Le déterminant d'une matrice ne change pas lorsqu'on ajoute à l'un de ses vecteurs-colonnes une combinaison linéaire des autres vecteurs-colonnes ; il est nul si l'un des vecteurs-colonnes est combinaison des autres.*

DÉMONSTRATION. On note $M = (m_{ij})$ la matrice dont les colonnes sont les vecteurs $x_1, \ldots, x_n$.
a) On a

$$\begin{aligned}
\det(x_{\rho(1)}, \ldots, x_{\rho(n)}) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma(\rho(j))} \\
&= \varepsilon(\rho) \sum_{\sigma \in S_n} \varepsilon(\rho)\varepsilon(\sigma) \prod_{j=1}^{n} m_{j\sigma\rho(j)} \quad \text{car} \quad \varepsilon(\rho)^2 = 1 \\
&= \varepsilon(\rho) \sum_{\sigma\rho \in S_n} \varepsilon(\sigma\rho) \prod_{j=1}^{n} m_{j\sigma\rho(j)} \quad \text{car} \quad \varepsilon(\sigma\rho) = \varepsilon(\sigma)\varepsilon(\rho) \\
&= \varepsilon(\rho) \sum_{\tau \in S_n} \varepsilon(\tau) \prod_{j=1}^{n} m_{j\tau(j)} \quad \text{en posant } \tau = \sigma\rho \\
&= \varepsilon(\rho)\det(x_1, \ldots, x_n)
\end{aligned}$$

En particulier si on note τ la transposition $(i\ j)$ alors $\varepsilon(\tau) = -1$ et par conséquent on a

$$\det(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) = \det(x_1, \ldots, x_{\tau(i)}, \ldots, x_{\tau(j)}, \ldots, x_n) = -\det(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n).$$

b) On a d'après la propriété précédente

$$\begin{aligned}
\det(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) &= -\det(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) \\
&= -\det(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \quad \text{car } x_i = x_j
\end{aligned}$$

et donc ce déterminant est nul.
c) Soit

$$x_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix} \text{ et } y_k = \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix},$$

alors en utilisant la deuxième formule du déterminant

$$\begin{aligned}
\det(x_1, \ldots, \lambda x_k + \mu y_k, \ldots, x_n) &= \sum_{\sigma \in S_n} \varepsilon(\sigma)(\lambda a_{\sigma(k)k} + \mu b_{\sigma(k)k}) \prod_{j=1, j\neq k}^{n} m_{\sigma(j)j} \\
&= \lambda \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(k)k} \prod_{j=1, j\neq k}^{n} m_{\sigma(j)j} + \mu \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(k)k} \prod_{j=1, j\neq k}^{n} m_{\sigma(j)j} \\
&= \lambda\det(x_1 \ldots, x_k, \ldots, x_n) + \mu\det(x_1 \ldots, y_k, \ldots, x_n)
\end{aligned}$$

d) Par la linéarité du déterminant par rapport à chaque variable (propriété précédente) on a

$$\det(x_1, \ldots, x_k + \sum_{l=1, l\neq k}^{n} \lambda_l x_l, \ldots, x_n) = \det(x_1, \ldots, x_k, \ldots, x_n) + \sum_{l=1, l\neq k}^{n} \lambda_l \det(x_1, \ldots, x_l, \ldots, x_n).$$

Or tous les déterminants de la somme sont nuls d'après la propriété **b)**, car ils ont deux colonnes identiques. Ceci montre le résultat voulu. En particulier, si un vecteur est égal à une combinaison linéaire des autres vecteurs, on ne change pas le déterminant si on ajoute au vecteur l'opposé de la combinaison linéaire. Donc on est ramené à calculer le déterminant d'une matrice dont une colonne est nulle et par conséquent ce déterminant est nul. □

Le prochain résultat ètablit le caractère multiplicatif du dèterminant.

Théorème 138.

Soient M et N deux matrices carrées d'ordre n, alors

$$\det(MN) = \det(M)\det(N).$$

Démonstration. Notons par $M_1, \dots, M_n$ les colonnes de M de sorte que celles du produit MN sont $\sum_{i=1}^n n_{i,1}M_i, \dots, \sum_{i=1}^n n_{i,n}M_i$. En utilisant la propriété **c)** de Proposition 137 on obtient

$$\begin{aligned}\det(MN) &= \det(\sum_{i_1=1}^{n} n_{i_1,1}M_{i_1}, \dots, \sum_{i_n=1}^{n} n_{i_n,n}M_{i_n}) \\ &= \sum_{1\leqslant i_1,\dots,i_n\leqslant n} n_{i_1,1}\dots n_{i_n,n}\det(M_{i_1}, \dots, M_{i_n}) \\ &= \sum_{1\leqslant i_1,\dots,i_n\leqslant n} n_{i_1,1}\dots n_{i_n,n}\det(M_{i_1}, \dots, M_{i_n}),\end{aligned}$$

et grâce à la propriété **b)** de Proposition 137 on peut restreindre la somme prècèdente aux indices $i_1, \dots, i_n$ distincts deux à deux. Notons

$$E := \{(i_1, \dots, i_n)/ \quad 1 \leqslant i_1, \dots, i_n \leqslant n, \quad \text{pour tout } 1 \leq j \neq k \leq n, \quad i_j \neq i_k\}.$$

Cet ensemble E est en bijection avec S_n, il suffit d'associer à un n-uplet $(i_1, \dots, i_n)$ de E la permutation σ de S_n telle que pour tout j entre 1 et n, $\sigma(j) = i_j$. Ainsi

$$\begin{aligned}\det(MN) &= \sum_{\sigma\in S_n} n_{\sigma(1),1}\dots n_{\sigma(n),n}\det(M_{\sigma(1)}, \dots, M_{\sigma(n)}) \\ &= \sum_{\sigma\in S_n} n_{\sigma(1),1}\dots n_{\sigma(n),n}\,\varepsilon(\sigma)\det(M_1, \dots, M_n) \\ &= \Big(\sum_{\sigma\in S_n} \varepsilon(\sigma)\, n_{\sigma(1),1}\dots n_{\sigma(n),n}\Big)\det M \\ &= \det(N)\det(M),\end{aligned}$$

grâce à la propriété **a)** de Proposition 137 et à Proposition 128. □

Remarque 139. *Il n'est pas inutile de répéter ici que le déterminant n'est pas linéaire ! On a en général*

$$\det(M + N) \neq \det(M) + \det(N).$$

On peut prendre par exemple le cas des matrices carrées :

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \text{ and } N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Corollaire 140. *Soient M et N deux matrices carrées d'ordre n, alors*

$$\det(MN) = \det(NM)$$

même si $MN \neq NM$.

Démonstration. D'après le théorème précédent on a :

$$\det(MN) = \det(M)\det(N) = \det(N)\det(M) = \det(NM).$$

□

Théorème 141.

Soit M une matrice carrée d'ordre n, alors

$$M \text{ est inversible} \iff \det(M) \neq 0,$$

et dans ce cas

$$\det(M^{-1}) = (\det(M))^{-1}.$$

Démonstration. On a

$$M \text{ est inversible } \iff MM^{-1} = \mathrm{I}_n \Rightarrow \det(M)\det(M^{-1}) = \det(MM^{-1}) = \det(\mathrm{I}_n) = 1,$$

et donc on en déduit que $\det(M) \neq 0$.

Pour la réciproque on va démontrer la contraposée. Si la matrice M n'est pas inversible, en particulier, comme c'est une matrice carrée, cela implique que son image n'est pas $\mathbb{R}^n$ tout entier, or cette image est l'espace vectoriel engendré par les vecteurs $Me_1, \ldots, Me_n$, où $e_1, \ldots, e_n$ sont les vecteurs de la base canonique. On se convainc rapidement que pour tout i, Me_i est la i ème colonne de M. Ainsi les vecteurs-colonnes de M sont liés, donc l'un des vecteurs est combinaison linéaire des autres et donc $\det(M) = 0$. Ainsi on en déduit $\det(M) \neq 0 \Rightarrow M$ est inversible. □

II.4. Determinant de matrices semblables et determinant d'un endomorphisme. Nous allons maintenant caractériser les déterminants de matrices qui sont semblables. Rappelons que deux matrices carrées d'ordre n, M et N, sont *semblables* s'il existe une matrice inversible P telle que $N = P^{-1}MP$. Pour que deux matrices soient semblables, il faut et il suffit qu'elles soient les matrices du même endomorphisme d'un espace vectoriel E dans deux bases.

Proposition 142. *Des matrices semblables ont même déterminant.*

Démonstration. On a

$$\det(N) = \det(P^{-1}MP) = \det(P^{-1})\det(M)\det(P) = \frac{1}{\det(P)}\det(M)\det(P) = \det(M).$$

□

Nous pouvons ainsi définir le déterminant d'un endomorphisme.

Théorème 143.

Soit E un K-espace vectoriel de dimension n et f un endomorphisme de E. Alors le scalaire $\det_{\mathcal{B}}(f(e_1), \ldots, f(e_n))$ ne dépend pas de la base $\mathcal{B} = (e_1, \ldots, e_n)$ choisie. On l'appelle déterminant de l'endomorphisme f *et on le note* $\det(f)$.

Démonstration. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ deux bases de E. On note M et N les matrices de l'endomorphisme f de E dans ces deux bases et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors on a $N = P^{-1}MP$ et d'après la proposition précédente $\det(M) = \det(N)$, d'où le résultat. □

Théorème 144.

Soient f et g deux endomorphismes d'un K-espace vectoriel E. Alors les propriétés suivantes sont vérifiées :

a) $\det(f \circ g) = \det(f)\det(g)$ *;*

b) f est bijectif si et seulement si $\det(f) \neq 0$ *et alors on a* $\det(f^{-1}) = \dfrac{1}{\det(f)}$.

Démonstration. Ces résultats s'obtiennent facilement lorsqu'on passe à l'écriture matricielle. □

II.5. Développement par rapport à une ligne ou une colonne.

Dans la pratique, il est extrêmement rare que l'on calcule un déterminant par la formule directe, qui est bien trop compliquée. On a recours à des astuces. Nous allons donner une formule qui permet effectivement de calculer le déterminant d'une matrice carrée d'ordre n.

Définition 145. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée d'ordre n. On note M_{ij} la matrice d'ordre $n-1$, obtenue à partir de M en supprimant la i ième ligne et la j ième colonne (tous les autres coefficients restent dans le même ordre). On appelle* mineur *de m_{ij} le scalaire* $\det(M_{ij})$ *et* cofacteur *de m_{ij} le scalaire* $\Delta_{ij} = (-1)^{i+j}\det(M_{ij})$.

Théorème 146.

Avec les notations de la définition précédente on a :
a) développement par rapport à la colonne j, $j = 1, \ldots, n$,

$$\det(M) = (-1)^{1+j}\, m_{1j}\det(M_{1j}) + \cdots + (-1)^{n+j}\, m_{nj}\det(M_{nj}) = m_{1j}\,\Delta_{1j} + \cdots + m_{nj}\,\Delta_{nj};$$

b) développement par rapport à la ligne i, $i = 1, \ldots, n$,

$$\det(M) = (-1)^{i+1}\, m_{i1}\det(M_{i1}) + \cdots + (-1)^{i+n}\, m_{in}\det(M_{in}) = m_{i1}\,\Delta_{i1} + \cdots + m_{in}\,\Delta_{in}.$$

Démonstration. Démontrons la première formule (la seconde se démontre de la même façon ou s'obtient en considérant la transposée de M). Pour une colonne de M, on peut écrire

$$\begin{pmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{nj} \end{pmatrix} = \begin{pmatrix} m_{1j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ m_{1j} \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{nj} \end{pmatrix}.$$

Comme le déterminant est $n-$linéaire, en particulier il est linéaire en la j ième colonne. Si on note N_{ij} la matrice obtenue à partir de M en remplaçant tous les éléments de la j ième colonne, sauf celui dans la i ième ligne, par 0, on obtient

$$\det(M) = \det(N_{1j}) + \cdots + \det(N_{nj}).$$

Si l'on fait passer la j ième colonne de N_{ij} à la première place, sans changer l'ordre des autres, on effectue $j-1$ transpositions de colonnes, donc le déterminant est multiplié par $(-1)^{j-1}$; de même, si l'on fait passer la i ième ligne de N_{ij} à la première place (sans changer l'ordre des autres) le déterminant est multiplié par $(-1)^{i-1}$. On a donc

$$\det(N_{ij}) = (-1)^{(i-1)+(j-1)} \begin{vmatrix} m_{ij} & m_{i1} & \cdots & m_{i\,j-1} & m_{i\,j+1} & \cdots & m_{in} \\ 0 & & & & & & \\ \vdots & & & M_{ij} & & & \\ 0 & & & & & & \end{vmatrix} = (-1)^{i+j}\, m_{ij}\det(M_{ij})$$

d'après la Proposition 131. On obtient donc le a). □

Exemple 147. *Il est avantageux de développer selon une ligne ou une colonne où il y a beaucoup de zéros :*

$$\begin{vmatrix}1&2&3&4\\4&0&5&6\\0&0&3&0\\1&6&4&2\end{vmatrix} = (-1)^{3+1}.0.\begin{vmatrix}2&3&4\\0&5&6\\6&4&2\end{vmatrix} + (-1)^{3+2}.0.\begin{vmatrix}1&3&4\\4&5&6\\1&4&2\end{vmatrix} + (-1)^{3+3}.3.\begin{vmatrix}1&2&4\\4&0&6\\1&6&2\end{vmatrix} + (-1)^{3+4}.0.\begin{vmatrix}1&2&3\\4&0&5\\1&6&4\end{vmatrix}$$

$$= 3.\begin{vmatrix}1&2&4\\4&0&6\\1&6&2\end{vmatrix} = 3.\begin{vmatrix}1&2&4\\4&0&6\\-2&0&-10\end{vmatrix} \quad L_3 \leftarrow L_3 - 3L_1$$

$$= 3.(-1)^{1+2}.2.\begin{vmatrix}4&6\\-2&-10\end{vmatrix} = -6(4.(-10) - 6.(-2)) = 168.$$

II.6. Calcul de l'inverse d'une matrice.

Nous verrons dans la suite que les déterminants vont nous être très utiles pour calculer le polynôme caractéristique d'un endomorphisme ou d'une matrice. Ici allons nous donner deux autres applications en commençant par le calcul de l'inverse d'une matrice.

Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On peut lui associer n^2 cofacteurs Δ_{ij} pour $1 \leq i, j \leq n$. Il est donc possible de construire une matrice carrée d'ordre n dont les coefficients sont les cofacteurs de M, c'est la matrice des cofacteurs dont la définition est donnée ci-dessous.

Définition 148. *Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On appelle* comatrice *de M la matrice notée* $\mathrm{Com}(M)$ *telle que le coefficient de la i ième ligne et de la j ième colonne soit exactement le cofacteur Δ_{ij} de M.*

Nous donnons maintenant une relation entre une matrice et sa comatrice, ainsi qu'une expression de l'inverse d'une matrice.

Théorème 149.

Soit $M = (m_{ij})_{1\leq i,j\leq n}$ une matrice carrée. On a

$$M\,{}^t\mathrm{Com}(M) = \det(M) I_n \text{ et } {}^t\mathrm{Com}(M)\,M = \det(M) I_n.$$

De plus si M est inversible ($\det(M) \neq 0$) *alors on a*

$$M^{-1} = \frac{1}{\det(M)}\,{}^t\mathrm{Com}(M).$$

Démonstration. Le terme d'indice (i, j) du produit $M\,{}^t\mathrm{Com}(M)$ est $\sum_{k=1}^{n} m_{ik}\Delta_{jk}$ (on multiplie par la transposée). Distinguons deux cas :

- Si $j = i$ on reconnaît la deuxième formule du théorème 146 et donc on obtient $\det(M)$.
- Sinon c'est le développement suivant la j ième ligne, du déterminant de la matrice obtenue à partir de M en ayant remplacé la j ième ligne par la i ième. Cette matrice a deux lignes égales, donc son déterminant est nul. On obtient donc $M\,{}^t\mathrm{Com}(M) = \det(M) I_n$.

L'autre égalité s'obtient de la même manière en utilisant le développement suivant les colonnes.

Pour le dernier résultat, lorsque M est inversible il suffit de diviser tous les termes dans la première formule par $\det(M)$. □

Exemple 150. *Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible c'est-à-dire telle que* $\det(M) = ad - bc \neq 0$. *Alors l'inverse de M est donné par*

$$M^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

II.7. Système de Cramer.

On considère ici des systèmes linéaires qui ont un nombre d'équations égal au nombre d'inconnues.

Définition 151. *Soit n un entier naturel non nul. On appelle* système d'équations linéaires *de n équations à n inconnues sur le corps K, le système :*

$$(S)\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

où les a_{ij} et b_i sont des éléments de K pour $i = 1, \ldots, n$ et $j = 1, \ldots, n$.

On peut écrire matriciellement ce système de la manière suivante. Soit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

la matrice du système. On pose encore

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Alors le système est équivalent à l'équation matricielle $AX = B$.

Théorème 152.

On dit qu'un système linéaire de n équations à n inconnues est de Cramer*, si sa matrice est inversible. Alors le système a une unique solution qui est*

$$X = A^{-1}B.$$

Démonstration. Le résultat est immédiat. □

Proposition 153. *Soit (S) un système de Cramer, d'équation matricielle $AX = B$. Pour i, $1 \leq i \leq n$, on définit la matrice A_i obtenue en remplaçant dans A la i ième colonne par le second membre B. Alors la solution du système est donnée par : pour tout i compris entre 1 et n*

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

Démonstration. Comme le système est de Cramer, il admet une unique solution. Si on note $C_1, \ldots, C_n$ les vecteurs colonnes de la matrice A du système, on peut écrire

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = B.$$

Maintenant, si A_i est la matrice obtenue en remplaçant dans A la i ième colonne par le second membre B, on a

$$\begin{aligned} \det(A_i) &= \det(C_1, \ldots, C_{i-1}, B, C_{i+1}, \ldots, C_n) \\ &= \det(C_1, \ldots, C_{i-1}, x_1C_1 + x_2C_2 + \cdots + x_nC_n, C_{i+1}, \ldots, C_n) \\ &= \sum_{k=1}^{n} x_k \det(C_1, \ldots, C_{i-1}, C_k, C_{i+1}, \ldots, C_n) \\ &= x_i \det(C_1, \ldots, C_{i-1}, C_i, C_{i+1}, \ldots, C_n) = x_i \det(A) \end{aligned}$$

car l'application déterminant est une forme n-linéaire alternée. Comme $\det(A) \neq 0$, ceci montre le résultat. □

Exemple 154. *On souhaite résoudre le système :*

$$(S)\begin{cases} 2x_1 + 3x_2 + x_3 = 9 \\ x_1 + 2x_2 + 3x_3 = 6 \\ 3x_1 + x_2 + 2x_3 = 8 \end{cases}.$$

La matrice du système est $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. *Son déterminant est égal à* 18, *c'est donc un système de Cramer. La solution est donc*

$$\begin{aligned} x_1 &= \frac{1}{18}\begin{vmatrix} 9 & 3 & 1 \\ 6 & 2 & 3 \\ 8 & 1 & 2 \end{vmatrix} = \frac{35}{18}, \\ x_2 &= \frac{1}{18}\begin{vmatrix} 2 & 9 & 1 \\ 1 & 6 & 3 \\ 3 & 8 & 2 \end{vmatrix} = \frac{29}{18}, \\ x_3 &= \frac{1}{18}\begin{vmatrix} 2 & 3 & 9 \\ 1 & 2 & 6 \\ 3 & 1 & 8 \end{vmatrix} = \frac{5}{18}. \end{aligned}$$

III. Diagonalisation

III.1. Valeur propre et vecteur propre.

Définition 1. *Soit $f \in \mathcal{L}(E)$. On dit que $\lambda \in K$ est* valeur propre *de f s'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$; x est alors appelé* vecteur propre *de f associé à la valeur propre λ.*

Remarque 2. *Tous les multiples non nuls d'un vecteur propre de f sont encore des vecteurs propres de f pour la même valeur propre. L'ensemble des valeurs propres d'un endomorphisme f s'appelle le spectre* de f et est noté $\mathrm{Sp}(f)$.

Exemple 3. *Si f est une homothétie d'un espace vectoriel E, $f = a\mathrm{Id}_E$, alors tout vecteur non nul est un vecteur propre associé à la valeur propre a.*

Exemple 4. *Comme on l'a vu dans l'introduction il existe des endomorphismes qui n'ont pas de valeur propre ni de vecteur propre ; par exemple les rotations d'angle différent de* $0 \mod \pi$ *dans le plan réel.*

Dans le théorème suivant nous caractérisons de façon plus précise les valeurs propres d'un endomorphisme.

Théorème 5.

Soient $f \in \mathcal{L}(E)$ et λ un scalaire. Les assertions suivantes sont équivalentes :
(i) *λ est valeur propre de f ;*
(ii) *l'endomorphisme $f - \lambda\mathrm{Id}_E$* n'est pas injectif, *i.e. son* noyau *vérifie*

$$\mathrm{Ker}(f - \lambda\mathrm{Id}_E) = \{x \in E,\ (f - \lambda\mathrm{Id}_E)(x) = 0\} \neq \{0\};$$

(iii) $\det(f - \lambda\mathrm{Id}_E) = 0$ *;*
(iv) $\det(M - \lambda\mathrm{I}_n) = 0$ *où M est la matrice de f dans n'importe quelle base de E.*

Démonstration. Pour que λ soit valeur propre de f il faut et il suffit qu'il existe un vecteur non nul x de E tel que $f(x) = \lambda x$, c'est-à-dire que l'on ait $(f - \lambda \mathrm{Id}_E)(x) = 0$, ou encore que le noyau $\mathrm{Ker}(f - \lambda \mathrm{Id}_E) \neq \{0\}$. Ceci entraîne l'équivalence de (i) et (ii). Pour que l'endomorphisme $f - \lambda \mathrm{Id}_E$ de l'espace vectoriel *de dimension finie* E ne soit pas injectif il faut et il suffit qu'il ne soit pas bijectif, c'est-à-dire que son déterminant soit nul, d'où l'équivalence de (ii) et (iii). Enfin par définition du déterminant d'un endomorphisme, (iii) et (iv) sont équivalentes. □

Ce qui précède montre que l'ensemble des vecteurs propres associés à une valeur propre λ auquel on ajoute le vecteur nul est exactement $\mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.

Définition 6. *Soit λ une valeur propre d'un endomorphisme f. On appelle* sous-espace propre *associé à λ, le sous-espace vectoriel de E défini par $E_\lambda = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$.*

Remarque 7. *C'est en cherchant le noyau de l'application $f - \lambda \mathrm{Id}_E$ que l'on détermine les vecteurs propres associés à la valeur propre λ.*

III.2. Polynôme caractéristique.

Définition 8. *Le* polynôme caractéristique *de $f \in \mathcal{L}(E)$ est défini par $\chi_f(X) = \det(f - X\mathrm{Id}_E)$.*

Si E est un K-espace vectoriel alors $\chi_f(X) \in K[X]$. De plus, si M est la matrice de f dans une base quelconque $\mathcal{B}$ de E, alors

$$\chi_f(X) = \det(M - X\mathrm{I}_n).$$

Théorème 9.

Les valeurs propres d'un endomorphisme f sur un K-espace vectoriel E sont exactement les racines de son polynôme caractéristique qui sont dans K.

Démonstration. On a les équivalences suivantes :

$$\begin{aligned} \lambda \in K \text{ est valeur propre de } f &\iff (f - \lambda \mathrm{Id}_E) \text{ est non injectif} \iff \det(f - \lambda \mathrm{Id}_E) = 0 \\ &\iff \chi_f(\lambda) = 0 \iff \lambda \text{ est une racine de } \chi_f \text{ dans } K \end{aligned}$$

□

Soit M une matrice d'ordre n à coefficients dans K. Soit X une indéterminée, alors on peut écrire :

$$M - X\mathrm{I}_n = \begin{pmatrix} m_{11} - X & m_{12} & \cdots & m_{1n} \\ m_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & m_{(n-1)n} \\ m_{n1} & \cdots & m_{n(n-1)} & m_{nn} - X \end{pmatrix}$$

Le déterminant de cette matrice est une somme de produits de coefficients de la matrice, c'est donc bien un polynôme en X à coefficients dans K. Chacun des produits comporte n termes, qui sont des coefficients de chacune des colonnes ; ce sont donc des polynômes de degré au plus n. De plus un seul produit est de degré n, c'est le produit de tous les coefficients diagonaux (il correspond à la permutation $\sigma = \mathrm{Id}$). Comme la signature de l'identité est égale à $+1$, le terme de plus haut degré est $(-1)^n X^n$. On a donc montré le résultat suivant.

Proposition 10. *Le polynôme caractéristique d'une matrice d'ordre n (ou d'un endomorphisme d'un espace vectoriel de dimension n) est un polynome de degré n dont le terme dominant est $(-1)^n$.*

Comme un polynôme de degré n a au plus n racines on obtient :

Corollaire 11. *En dimension n un endomorphisme (ou une matrice d'ordre n) a au plus n valeurs propres distinctes.*

Exemple 12. *Le polynôme caractéristique de la matrice*

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

est

$$\chi_M(X) = \begin{vmatrix} -X & 1 \\ 1 & -X \end{vmatrix} = X^2 - 1.$$

Les valeurs propres sont ± 1*. Les sous-espaces propres associés à ces valeurs propres se déterminent alors de la manière suivante :*

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_1 \iff \begin{cases} -x + y = 0 \\ x - y = 0 \end{cases} \iff \{x = y \iff E_1 = < \begin{pmatrix} 1 \\ 1 \end{pmatrix} > .$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \in E_{-1} \iff \begin{cases} x + y = 0 \\ x + y = 0 \end{cases} \iff \{y = -x \iff E_1 = < \begin{pmatrix} 1 \\ -1 \end{pmatrix} > .$$

Exemple 13. *La rotation plane* R_θ *a pour matrice dans la base canonique*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Son polynôme caractéristique est

$$\chi_{R_\theta}(X) = \begin{vmatrix} \cos\theta - X & -\sin\theta \\ \sin\theta & \cos\theta - X \end{vmatrix} = (\cos\theta - X)^2 + (\sin\theta)^2 = X^2 - 2X\cos\theta + 1.$$

Ce polynôme n'a pas de racines réelles si $\sin\theta$ *est non nul. Par contre dans* $\mathbb{C}$ *il a deux racines* $e^{\pm i\theta}$*, qui sont donc les valeurs propres de* R_θ*. On peut vérifier que les vecteurs propres associés sont*

$$\begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

Exemple 14. *Soit* $M = (m_{ij})_{1 \le i,j \le n}$ *une matrice triangulaire. Alors* $M - X\mathrm{I}_n$ *est aussi une matrice triangulaire et le polynôme caractéristique (déterminant d'une matrice triangulaire) est donc le produit des coefficients diagonaux i.e.* $\chi_M(X) = (m_{11} - X)\cdots(m_{nn} - X)$*. Les valeurs propres de* M *sont donc les coefficents diagonaux de* M*. En particulier ce résultat est vrai pour une matrice diagonale.*

Définition 15. *Soit* f *un endomorphisme d'un* K*-espace vectoriel* E*.*

(1) L'ordre de multiplicité d'une valeur propre λ *de* f *est l'ordre de multiplicité de la racine* λ *dans le polynôme caractéristique de* f*.*

(2) Un polynôme de $K[X]$ *est dit* scindé *si toutes ses racines sont dans* K *autrement dit s'il se décompose dans* $K[X]$ *comme produit de polynômes de degré 1.*

(3) Par extension on dira qu'un endomorphisme est scindé si son polynôme caractéristique l'est.

On peut remarquer que si un endomorphisme f d'un espace vectoriel E de dimension n, est scindé, son polynôme caractéristique a la forme suivante :

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où r, $1 \le r \le n$, représente le nombre de valeurs propres distinctes, les $(\lambda_i)_{1 \le i \le r}$ sont les différentes valeurs propres et les $(\alpha_i)_{1 \le i \le r}$ sont leurs ordres de multiplicité respectifs. On a de plus $\sum_{i=1}^{r} \alpha_i = n$.

III.3. Étude des sous-espaces propres.

Théorème 16.

Soit $f \in \mathcal{L}(E)$. On considère $\{\lambda_1, \dots, \lambda_r\}$ l'ensemble des valeurs propres deux à deux distinctes de f. Alors les sous-espaces propres $(E_{\lambda_i})_{1\leq i\leq r}$ sont en somme directe, *i.e. $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.*

Rappelons que r sous-espaces vectoriels $(E_{\lambda_i})_{1\leq i\leq r}$ sont dits en somme directe si et seulement si l'une deux propriétés équivalentes suivantes est vérifiée :

(**a**) $x_1 + x_2 + \cdots + x_r = 0$, avec pour $i = 1, \dots, r$, $x_i \in E_{\lambda_i} \Rightarrow \forall\ 1 \leq i \leq r,\ x_i = 0$;

(**b**) pour $1 \leq k \leq r$, $E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}$.

Démonstration. On va établir ce résultat par récurrence sur r le nombre de valeurs propres distinctes. Si $r = 2$, alors on a $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. En effet soit $x \in E_{\lambda_1} \cap E_{\lambda_2}$, alors $f(x) = \lambda_1 x$ et $f(x) = \lambda_2 x$. On obtient donc $(\lambda_1 - \lambda_2)x = 0$ c'est-à-dire $x = 0$ car $\lambda_1 \neq \lambda_2$. Un vecteur propre est donc toujours associé qu'à une seule valeur propre.

Supposons maintenant la propriété vraie jusqu'au rang $k-1$, c'est-à-dire que pour tout l entre 2 et $k-1$,

$$E_{\lambda_l} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{l-1}}) = \{0\}.$$

Montrons alors que :

$$E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}.$$

Soit

$$x \in E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}).$$

On a $x = x_1 + \cdots + x_{k-1}$ avec pour $1 \leq i \leq k-1$, $x_i \in E_{\lambda_i}$. En prenant l'image par f de cet élément, on obtient :

$$f(x) = \lambda_k x = \lambda_k(x_1 + \cdots + x_{k-1}),$$

et d'autre part

$$f(x) = f(x_1) + \cdots + f(x_{k-1}) = \lambda_1 x_1 + \cdots + \lambda_{k-1} x_{k-1}.$$

Ceci entraîne que

$$(\lambda_k - \lambda_1)x_1 + \cdots + (\lambda_k - \lambda_{k-1})x_{k-1} = 0.$$

Or par hypothèse de récurrence $E_{\lambda_1}, \dots, E_{\lambda_{k-1}}$ sont en somme directe, et comme les valeurs propres sont deux à deux distinctes, on obtient $x_i = 0$ pour $1 \leq i \leq k-1$, c'est-à-dire $E_{\lambda_k} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{k-1}}) = \{0\}$. □

Théorème 17.

Soient $f \in \mathcal{L}(E)$ et λ une valeur propre de f. Alors la dimension du sous-espace propre associé à la valeur propre λ est inférieure ou égale à la multiplicité de la valeur propre λ, c'est-à-dire : $\dim(E_\lambda) \leq \alpha_\lambda$.
En particulier, si λ est une valeur propre simple (multiplicité égale à 1) alors $\dim(E_\lambda) = 1$

Démonstration. On considère $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ une base de E telle que $(e_1, \dots, e_k)$ soit une base de E_λ. Dans cette base la matrice de f s'écrit

$$M = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & \mathcal{M}_1 \\ & & \lambda & \\ \hline 0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & \mathcal{M}_2 \\ 0 & \cdots & 0 & \end{array}\right)$$

Le polynôme caractéristique de f est donc

$$\chi_f(X) = \det(M - X\mathrm{I}_n) = (\lambda - X)^k \det(\mathcal{M}_2 - X\mathrm{I}_{n-k}).$$

Ainsi λ est racine de χ_f d'ordre supérieur ou égal à $k = \dim(E_\lambda)$. Pour le second point un sous-espace propre contient au-moins un vecteur propre (vecteur non nul) donc on a $1 \leq \dim(E_\lambda)$. Si on applique alors le premier point avec $\alpha_\lambda = 1$, on obtient le résultat. □

III.4. Endomorphismes diagonalisables.

Définition 18. *On dit que $f \in \mathcal{L}(E)$ est* diagonalisable *s'il existe une base de E constituée de vecteurs propres.*

Remarque 19. *Dans une telle base la matrice de f s'écrit*

$$M = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Théorème 20.

Soit $f \in \mathcal{L}(E)$, dont $\{\lambda_1, \ldots, \lambda_r\}$ est l'ensemble des valeurs propres deux à deux distinctes. f est diagonalisable si et seulement si ses sous-espaces propres vérifient : $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$.

Démonstration. Supposons que f soit diagonalisable. Quitte à réordonner les vecteurs, la base de vecteurs propres de f est de la forme

$$(e_{1,\lambda_1}, \ldots, e_{\alpha_1,\lambda_1}, \ldots, e_{1,\lambda_r}, \ldots, e_{\alpha_r,\lambda_r})$$

où pour $1 \leq i \leq r$ et pour $1 \leq j \leq \alpha_i$, e_{j,λ_i} est un vecteur propre associé à la valeur propre λ_i. Si $\alpha_i < \dim(E_{\lambda_i}) = \beta_i$ alors E_{λ_i} aurait une base de la forme $(e_{1,\lambda_i}, \ldots, e_{\beta_i,\lambda_i})$. De plus, d'après le théorème 9,

$$(e_{1,\lambda_1}, \ldots, e_{\alpha_1,\lambda_1}, \ldots, e_{1,\lambda_i}, \ldots, e_{\beta_i,\lambda_i}, \ldots, e_{1,\lambda_r}, \ldots, e_{\alpha_r,\lambda_r})$$

serait une famille libre. Ainsi on obtient une contradiction (car la famille libre aurait plus de vecteurs qu'une base). Donc pour tout $1 \leq i \leq r$, $\alpha_i = \dim(E_{\lambda_i})$, et comme les sous-espaces propres sont en somme directe, on aboutit au résultat.
Réciproquement, si les sous-espaces propres de f vérifient

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}.$$

Alors en notant une base de E_{λ_i}, $\mathcal{B}_i = (e_{1,\lambda_i}, \ldots, e_{\alpha_i,\lambda_i})$, on a $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E formée de vecteurs propres. □

Théorème 21.

Soit $f \in \mathcal{L}(E)$, avec E un $K-$espace vectoriel. f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :

- **(a)** *f est scindé, i.e. toutes ses valeurs propres sont dans K ;*
- **(b)** *Pour toute valeur propre λ de f on a $\dim(E_\lambda) = \alpha_\lambda$ qui est l'ordre de multiplicité de λ.*

DÉMONSTRATION. Si f est diagonalisable alors d'après le théorème 20 il existe une base de vecteurs propres, $(e_{1,\lambda_1}, \ldots, e_{\alpha_1,\lambda_1}, \ldots, e_{1,\lambda_r}, \ldots, e_{\alpha_r,\lambda_r})$, dans laquelle sa matrice s'écrit

$$M = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix}$$

Son polynôme caractéristique est égal à $(-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}$. Il est donc scindé, et pour $1 \leq i \leq r$, $\dim(E_{\lambda_i}) = \alpha_i$ = l'ordre de multiplicité de λ_i.
Réciproquement, pour tout sous-espace propre E_{λ_i} on a d'après la propriété (b), $\dim(E_{\lambda_i}) = \alpha_i$, et d'après (a),

$$\sum_{i=1}^{r} \alpha_i = n = \dim(E).$$

Donc comme les sous-espaces propres sont en somme directe, on obtient que $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

Corollaire 22. *Soit $f \in \mathcal{L}(E)$. Une condition suffisante pour que f soit diagonalisable est que f ait exactement n valeurs propres deux à deux distinctes.*

DÉMONSTRATION. Si f possède n valeurs propres deux à deux distinctes, f est scindé, et de plus tout sous-espace propre est de dimension 1. f est donc diagonalisable d'après le théorème 13. □

Attention ce n'est pas une condition nécessaire. En effet, les homothéties $\lambda \mathrm{Id}_E$, par exemple, sont diagonalisables mais n'ont qu'une seule valeur propre λ. Ce sont d'ailleurs les seuls endomorphismes diagonalisables qui ont une seule valeur propre.

III.5. Exemple de diagonalisation.
Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{pmatrix}.$$

Le polynôme caractéristique de f est alors $\chi_f(X) = (1 - X)^2(2 - X)$. f a donc une valeur propre double $\lambda = 1$, et une valeur propre simple $\lambda = 2$. Recherchons maintenant les sous-espaces propres. E_1 le sous-espace propre associé à la valeur propre $\lambda = 1$ est déterminé par :

$$(M - 1\mathrm{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 4x - 2y + 3z = 0 \\ 4x - 2y + 3z = 0 \end{cases} \iff \{\ 4x - 2y + 3z = 0$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 2, engendré par deux vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \text{ et } \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Le sous-espace propre E_2 associé à la valeur propre $\lambda = 2$ est déterminé par :

$$(M - 2\mathrm{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} -x & = & 0 \\ 4x - 3y + 3z & = & 0 \\ 4x - 2y + 2z & = & 0 \end{cases} \iff \begin{cases} x & = & 0 \\ y & = & z \end{cases}$$

C'est donc un sous-espace vectoriel de $\mathbb{R}^3$ de dimension 1, engendré par un vecteur u_3 de coordonnées dans la base canonique

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

La somme des dimensions des sous-espaces propres est égale à 3, c'est-à-dire à celle de l'espace vectoriel E (ici $\mathbb{R}^3$), f est donc diagonalisable et sa matrice dans la base (u_1, u_2, u_3) est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

IV. Trigonalisation

IV.1. Endomorphismes trigonalisables.

Définition 23. *Une matrice carrée M d'ordre n est dite* trigonalisable *si elle est semblable à une matrice triangulaire T, supérieure ou inférieure, c'est-à-dire s'il existe une matrice inversible P telle que $M = P^{-1}TP$.*

Définition 24. *Soit $f \in \mathcal{L}(E)$. f est dit* trigonalisable *s'il existe une base B de E dans laquelle sa matrice est triangulaire supérieure ou inférieure.*

On peut remarquer que tout endomorphisme (ou matrice) diagonalisable est trigonalisable. D'autre part, si $f \in \mathcal{L}(E)$ est trigonalisable cela signifie qu'il existe une base B de E dans laquelle la matrice de f s'écrit

$$M = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Son polynôme caractéristique est donc égal à $\chi_f(X) = (-1)^n \prod_{i=1}^{n}(X - a_{ii})$. Il est donc scindé et ses valeurs propres sont exactement les $(a_{ii})_{1 \leq i \leq n}$.

Théorème 25.

$f \in \mathcal{L}(E)$ est trigonalisable si et seulement si f est scindé.

Démonstration. Si f est trigonalisable, d'après la remarque ci-dessus, f est scindé. Réciproquement supposons que f soit scindé et montrons par récurrence sur la dimension n de E que f est trigonalisable. Si $n = 1$, la propriété est vraie. Supposons qu'elle soit vraie jusqu'au rang $n - 1$. Comme χ_f est scindé, il existe au moins une valeur propre λ. Soit alors u_1 un vecteur propre associé à cette valeur propre. On considère les vecteurs $(u_2, \ldots, u_n)$ de façon à ce que $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ soit une base de E. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & M_1 & \\ 0 & & & \end{pmatrix}$$

M_1 est la matrice d'un endomorphisme g égal à la composée de f avec la projection de E sur le sous-espace de E de dimension $n - 1$ engendré par $(u_2, \ldots, u_n)$. On a

$$\chi_f(X) = (\lambda - X)\det(M_1 - X\mathrm{I}_{n-1}) = (\lambda - X)\chi_g(X),$$

et comme χ_f est scindé, χ_g l'est aussi. Donc par hypothèse de récurrence, il existe une base $(u'_2, \dots, u'_n)$ de ce sous-espace dans laquelle la matrice de g est triangulaire supérieure. Ainsi dans la base $\mathcal{B}' = (u_1, u'_2, \dots, u'_n)$ de E la matrice de f est triangulaire supérieure. □

Corollaire 26. *Tout endomorphisme sur un espace vectoriel complexe est trigonalisable.*

Démonstration. Cela provient du fait que tout polynôme sur $\mathbb{C}[X]$ est scindé. □

IV.2. Exemple de trigonalisation.
On considère f l'endomorphisme de $E = \mathbb{R}^4$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}.$$

Le calcul du polynôme caractéristique donne $\chi_f(X) = \chi_M(X) = (1-X)^4$. Il y a donc une seule valeur propre $\lambda = 1$ d'ordre 4. On détermine maintenant le sous-espace propre E_1 associé à cette valeur propre.

$$(M - 1\mathrm{I}_4)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{rcl} -3y + 3t & = & 0 \\ -2x - 7y + 13t & = & 0 \\ -3y + 3t & = & 0 \\ -x - 4y + 7t & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcl} x & = & 3t \\ y & = & t \end{array} \right. .$$

L'unique sous-espace propre E_1 est donc de dimension 2 ($< 4 = \dim(\mathbb{R}^4)$ donc f n'est pas diagonalisable), il est engendré par les vecteurs u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

On complète (u_1, u_2) par (u_3, u_4) pour obtenir une base de $\mathbb{R}^4$. Ici on peut par exemple choisir $u_3 = e_1$ et $u_4 = e_2$. On a donc les relations suivantes :

$$e_1 = u_3,\ e_2 = u_4,\ e_3 = u_2 \text{ et } e_4 = u_1 - 3u_3 - u_4.$$

Ainsi :

$$\left\{ \begin{array}{rcl} f(u_3) & = & f(e_1) = e_1 - 2e_2 - e_4 = -u_1 + 4u_3 - u_4 \\ f(u_4) & = & f(e_2) = -3e_1 - 6e_2 - 3e_4 - 4e_4 = -4u_1 - 3u_2 + 9u_3 - 2u_4 \end{array} \right.$$

La matrice de f dans la base (u_1, u_2, u_3, u_4) s'écrit alors

$$\widetilde{M} = \begin{pmatrix} 1 & 0 & -1 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 4 & 9 \\ 0 & 0 & -1 & -2 \end{pmatrix}.$$

On considère la sous-matrice

$$M_1 = \begin{pmatrix} 4 & 9 \\ -1 & -2 \end{pmatrix}$$

qui est la matrice de la projection de f sur le sous-espace engendré par u_3 et u_4, dans la base canonique de ce sous-espace. On va maintenant trigonaliser cette matrice. Le polynôme caractéristique de cette matrice est

$$\chi_{M_1}(X) = (1-X)^2.$$

Cette matrice n'a qu'une seule valeur propre double qui est 1. Le sous-espace propre associé a cette valeur propre est déterminé par

$$\left\{ \begin{array}{rcl} 3x + 9y & = & 0 \\ -x - 3y & = & 0 \end{array} \right. \iff \{x = -3y\ .$$

Sa dimension est donc 1, et il est engendré par v_1 de coordonnées dans la base canonique du sous-espace

$$\begin{pmatrix}-3\\1\end{pmatrix}.$$

On le complète en une base du sous-espace avec un vecteur v_2 de coordonnées

$$\begin{pmatrix}0\\1\end{pmatrix}.$$

Les vecteurs correspondants dans l'espace $E=\mathbb{R}^4$ sont donc $v_1'=-3u_3+u_4$ et $v_2'=u_4$. Ainsi

$$\left\{\begin{array}{rcl} f(v_1') &=& -u_1-3u_2+v_1' \\ f(v_2') &=& -4u_1-3u_2-3v_1'+v_2' \end{array}\right. .$$

La matrice de f dans la base (u_1,u_2,v_1',v_2') s'écrit alors

$$\begin{pmatrix}1&0&-1&-4\\0&1&-3&-3\\0&0&1&-3\\0&0&0&1\end{pmatrix}.$$

V. Polynômes d'endomorphismes - Polynôme minimal

V.1. Polynômes d'endomorphismes.

Soit E un $K-$espace vectoriel de dimension n et f un endomorphisme de E. On introduit les notations suivantes :

$$f^0=\mathrm{Id}_E,\quad f^1=f,\quad f^2=f\circ f,\quad f^k=\overbrace{f\circ f\circ\cdots\circ f}^{k\text{ fois}}.$$

Plus généralement, si $Q(X)=a_0+a_1X+\cdots+a_kX^k$ est un polynôme de $K[X]$, alors on définit le *polynôme d'endomorphismes* $Q(f)$ par :

$$Q(f)=a_0\mathrm{Id}_E+a_1f+\cdots+a_kf^k.$$

Si M est la matrice de f dans une base $\mathcal{B}$ de E alors le *polynôme de matrices* $Q(M)$ défini par :

$$Q(M)=a_0\mathrm{I}_n+a_1M+\cdots+a_kM^k,$$

est la matrice de $Q(f)$ dans la base $\mathcal{B}$.

Proposition 27. *Pour tout endomorphisme f de E, il existe un polynôme non nul Q de $K[X]$ tel que $Q(f)=0$.*

Démonstration. E est un $K-$espace vectoriel de dimension n, donc $\mathcal{L}(E)$ est un $K-$espace vectoriel de dimension n^2. Par conséquent, les n^2+1 endomorphismes $\mathrm{Id}_E,\ f,\ f^2,\ldots,\ f^{n^2}$ sont liés. Il existe donc des coefficients $a_0,\ a_1,\ldots,\ a_{n^2}$ de K non tous nuls, tels que $a_0\mathrm{Id}_E+a_1f+\cdots+a_{n^2}f^{n^2}=0$. C'est-à-dire que le polynôme non nul

$$Q(X)=a_o+a_1X+\cdots+a_{n^2}X^{n^2}$$

de $K[X]$ vérifie $Q(f)=0$. □

V.2. Polynôme minimal.

Théorème 28.

Soient E un $K-$espace vectoriel et f un endomorphisme de E. L'ensemble $\mathfrak{I}=\{P\in K[X],\ P(f)=0\}$ des polynômes annulateurs de f est un idéal différent de $\{0\}$. Cet idéal est appelé idéal annulateur *de f. Le générateur unitaire de cet idéal s'appelle le* polynôme minimal *de f et est noté $m_f(X)$.*

DÉMONSTRATION. D'après la Proposition précédente cet ensemble contient un élément non nul. Donc $\mathfrak{I}$ est non vide et est différent de $\{0\}$. Soient P_1 et P_2 deux éléments de $\mathfrak{I}$. Alors $(P_1-P_2)(f) = P_1(f)-P_2(f) = 0$, c'est-à-dire que P_1-P_2 appartient à $\mathfrak{I}$. De plus si $P \in \mathfrak{I}$ et si $Q \in K[X]$ alors on a $(PQ)(f) = P(f)Q(f) = 0$. Ainsi $PQ \in \mathfrak{I}$. Ainsi $\mathfrak{I}$ est un idéal de $K[X]$. □

Remarque 29. *Le polynôme minimal d'un endomorphisme f est donc l'unique polynôme de plus bas degré et unitaire (coefficient dominant est égal à 1), vérifiant $m_f(f) = 0$. De plus si M est la matrice de f dans n'importe quelle base $\mathcal{B}$ de E, $m_M(X) = m_f(X)$ et donc $m_f(M) = 0$. Comme un générateur d'un idéal est de degré supérieur ou égal à 1, il en est de même pour le polynôme minimal .*

Exemple 30. *Soit f une homothétie d'un $K-$espace vectoriel E. On a $f = \lambda \mathrm{Id}_E$, avec $\lambda \in K$. Donc le polynôme $X - \lambda$ est un polynôme annulateur de f. Comme il est de degré 1 et unitaire, c'est le générateur unitaire de l'idéal annulateur de f, c'est-à-dire que c'est le polynôme minimal de f. Réciproquement, si le polynôme minimal de f est $X - \lambda$, alors on a $f - \lambda \mathrm{Id}_E = 0$, ou encore $f = \lambda \mathrm{Id}_E$. Par conséquent, f est une homothétie. Nous venons donc de montrer que le polynôme minimal d'un endomorphisme est de degré 1 si et seulement si cet endomorphisme est une homothétie.*

V.3. Théorème de Cayley-Hamilton.

Théorème 31.

Soient E un $K-$espace vectoriel et f un endomorphisme de E. Le polynôme minimal de f divise le polynôme caractéristique de f.

DÉMONSTRATION. D'après la remarque ci-dessus il suffit de montrer que $\chi_f(f) = 0$. Soit M la matrice de f dans une base $\mathcal{B}$ de E. C'est une matrice carrée d'ordre n. On pose $N = M - X\mathrm{I}_n$ et on note $\mathrm{Com}(N)$ sa comatrice. Les coefficients de $\mathrm{Com}(N)$ et donc de ${}^t\mathrm{Com}(N)$ sont des polynômes de degré inférieur ou égal à $n-1$. On peut donc écrire

$$ {}^t\mathrm{Com}(N) = \sum_{j=0}^{n-1} A_j X^j $$

où les A_j sont des matrices carrées à coefficients dans K. D'après le théorème 24 du chapitre précédent on a

$$ N\,{}^t\mathrm{Com}(N) = \det(N)\mathrm{I}_n = \chi_f(X)\mathrm{I}_n, $$

c'est-à-dire

$$ (M - X\mathrm{I}_n)\sum_{j=1}^{n-1} A_j X^j = \chi_f(X)\mathrm{I}_n. $$

En notant maintenant

$$ \chi_f(X) = \sum_{j=0}^{n} a_j X^j $$

(avec $a_n = (-1)^n$) on obtient en identifiant les coefficients de même degré :

$$ MA_0 = a_o\mathrm{I}_n,\ MA_1 - A_0 = a_1\mathrm{I}_n,\ldots,\ MA_{n-1} - A_{n-2} = a_{n-1}\mathrm{I}_n,\ -A_{n-1} = a_n\mathrm{I}_n. $$

On en déduit

$$ \chi_f(M) = \sum_{j=0}^{n} a_j M^j = MA_0 + M(MA_1 - A_0) + \cdots + M^{n-1}(MA_{n-1} - A_{n-2}) + M^n(-A_{n-1}) = 0, $$

car les termes se regroupent et s'annulent. □

Remarque 32. *Un énoncé équivalent de ce résultat est que χ_f appartient à l'idéal annulateur*

$$\mathfrak{I} = \{P \in K[X],\ P(f) = 0\},$$

ou encore que $\chi_f(f) = 0$. Comme le polynôme caractéristique est de degré n, on en déduit que le polynôme minimal est au plus de degré n.

Exemple 33. *Soit la matrice*

$$M = \begin{pmatrix} 0 & 3 \\ -2 & -5 \end{pmatrix}.$$

Le polynôme caractéristique de cette matrice est

$$\begin{aligned} \chi_M(X) &= \begin{vmatrix} -X & 3 \\ -2 & -X-5 \end{vmatrix} = \begin{vmatrix} -X-2 & -X-2 \\ -2 & -X-5 \end{vmatrix} = (-X-2)\begin{vmatrix} 1 & 1 \\ -2 & -X-5 \end{vmatrix} \\ &= -(X+2)\begin{vmatrix} 1 & 0 \\ -2 & -X-3 \end{vmatrix} = -(X+2)(-X-3) = (X+2)(X+3). \end{aligned}$$

Comme le polynôme minimal, divise le polynôme caractéristique, et qu'il est unitaire de degré au moins 1, on en déduit qu'il est égal à $X+2$, $X+3$ ou $(X+2)(X+3)$. Comme $M + 2\mathrm{I}_2 \neq 0$ et que $M + 3\mathrm{I}_2 \neq 0$, on obtient que $m_M(X) = (X+2)(X+3)$.

Proposition 34. *λ est valeur propre de f si et seulement si λ est une racine de son polynôme minimal.*

Démonstration. Si λ est racine de m_f alors comme m_f divise χ_f, λ est racine de χ_f, et donc d'après le théorème 9, λ est une valeur propre de f.
Réciproquement, soit λ une valeur propre de f dont on note x un vecteur propre associé. On effectue la division euclidienne de m_f par $X - \lambda$, et on obtient $m_f(X) = Q(X)(X-\lambda) + c$, où $\deg(c) \leq 0$, i.e. $c \in K$. On en déduit que

$$0 = m_f(f) = Q(f) \circ (f - \lambda \mathrm{Id}_E) + c\mathrm{Id}_E.$$

Si on applique cette relation au vecteur x, on trouve

$$0 = Q(f) \circ (f - \lambda \mathrm{Id}_E)(x) + c\mathrm{Id}_E(x) = Q(f)(f(x) - \lambda x) + cx = Q(f)(0) + cx = cx.$$

Or comme x n'est pas nul, ceci entraîne que $c = 0$, de sorte que le polynôme minimal s'écrit

$$m_f(X) = Q(X)(X - \lambda).$$

Cela signifie que λ est racine de m_f. □

Exemple 35. *Si on reprend l'exemple précédent, on a trouvé que le polynôme caractéristique de la matrice M est $\chi_M(X) = (X+2)(X+3)$. C'est-à-dire que les valeurs propres de M sont -2 et -3. Donc d'après ce qui précède, ce sont des racines du polynôme minimal. D'autre part comme il est unitaire et qu'il divise le polynôme caractéristique, on obtient directement que $m_M(X) = (X+2)(X+3)$.*

Exemple 36. *Soit la matrice*

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 2 \end{pmatrix}.$$

Comme la matrice est triangulaire, le polynôme caractéristique de cette matrice est $\chi_M(X) = (1-X)(1-X)(2-X)$. Les valeurs propres de M sont donc 1 et 2. Par définition du polynôme minimal et d'après le théorème de Cayley-Hamilton, on a $m_M(X) = (X-1)(X-2)$ ou $(X-1)^2(X-2)$, car toute valeur propre est racine de ce polynôme. Or lorsqu'on effectue le calcul suivant :

$$(M - \mathrm{I}_3)(M - 2\mathrm{I}_3) = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & -5 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -10 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

Donc le polynôme minimal est $m_M(X) = (X-1)^2(X-2)$.

V.4. Lemme de décomposition des noyaux.
Les résultats que l'on démontre ici reposent sur l'énoncé suivant.

Lemme 37. *Soit f un endomorphisme d'un $K-$espace vectoriel E et P un polynôme de $K[X]$.*
(a) *Le sous-espace vectoriel* $\mathrm{Ker}P(f)$ *est stable par f, c'est-à-dire*

$$f\left(\mathrm{Ker}P(f)\right) \subset \mathrm{Ker}P(f).$$

(b) *Si $P = P_1P_2$ avec P_1 et P_2 deux polynômes premiers entre eux, alors*

$$\mathrm{Ker}P(f) = \mathrm{Ker}P_1(f) \oplus \mathrm{Ker}P_2(f).$$

Démonstration. Remarquons d'abord que f et $P(f)$ commutent. Si x appartient au noyau de $P(f)$ on a $P(f)\big(f(x)\big) = f\big(P(f)(x)\big) = 0$, de sorte que $f(x)$ est encore dans le noyau de $P(f)$. Cela montre a).

Comme P_1 et P_2 sont premiers entre eux, d'après le théorème de Bézout il existe Q_1 et Q_2 deux polynômes de $K[X]$ tels que

$$Q_1P_1 + Q_2P_2 = 1.$$

En appliquant cette relation à f on obtient

$$Q_1(f)P_1(f) + Q_2(f)P_2(f) = \mathrm{Id}_E.$$

Soit $x \in \mathrm{Ker}P_1(f) \cap \mathrm{Ker}P_2(f)$, on a $P_1(f)(x) = P_2(f)(x) = 0$, et donc d'après l'égalité ci-dessus

$$x = \mathrm{Id}_E(x) = Q_1(f)P_1(f)(x) + Q_2(f)P_2(f)(x) = 0.$$

$\mathrm{Ker}P_1(f)$ et $\mathrm{Ker}P_2(f)$ sont donc en somme directe. D'autre part, pour tout x dans $\mathrm{Ker}P(f)$ on a :

$$x = \mathrm{Id}_E(x) = Q_1(f)P_1(f)(x) + Q_2(f)P_2(f)(x).$$

Or, comme

$$P_2(f)\big(Q_1(f)P_1(f)(x)\big) = Q_1(f)\big(P(f)(x)\big) = 0,$$

$Q_1(f)P_1(f)(x)$ est dans le noyau de $P_2(f)$; de même le vecteur $Q_2(f)P_2(f)(x)$ est dans le noyau de $P_1(f)$. Cela montre b). □

Corollaire 38. *Soit f un endomorphisme d'un $K-$espace vectoriel E et l polynômes $P_1, \ldots, P_l$ de $K[X]$ premiers entre eux deux à deux. On pose $P = P_1P_2 \ldots P_l$. Alors*

$$\mathrm{Ker}P(f) = \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_l(f).$$

Démonstration. On procède par récurrence sur l. Le cas $l = 1$ est trivial. Le cas $l = 2$ est exactement celui du Lemme 37. On suppose la propriété vraie au rang $l-1$ et on montre qu'elle l'est encore au rang l. Comme les polynômes sont premiers entre eux deux à deux P_l est premier avec $P_1P_2 \ldots P_{l-1}$. Donc par l'hypothèse de récurrence on a :

$$\begin{aligned}\mathrm{Ker}P(f) &= \mathrm{Ker}(P_1P_2 \ldots P_{l-1})(f) \oplus \mathrm{Ker}P_l(f) = \big(\mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f)\big) \oplus \mathrm{Ker}P_l(f) \\ &= \mathrm{Ker}P_1(f) \oplus \cdots \oplus \mathrm{Ker}P_{l-1}(f) \oplus \mathrm{Ker}P_l(f).\end{aligned}$$

□

V.5. Diagonalisation à l'aide du polynôme minimal.

Théorème 39.

Soient E un $K-$espace vectoriel et f un endomorphisme de E de valeurs propres distinctes $\{\lambda_1, \dots \lambda_r\}$. Les assertions suivantes sont équivalentes :
(i) *f est diagonalisable ;*
(ii) $m_f(X) = \prod_{i=1}^{r}(X - \lambda_i)$ *;*
(iii) *m_f est scindé et a racines simples ;*
(iv) *f admet un polynôme annulateur scindé à racines simples.*

Démonstration. $(i) \Rightarrow (ii)$: d'après la proposition 24 les valeurs propres sont les racines de m_f. Donc le polynome

$$M(X) = \prod_{i=1}^{r}(X - \lambda_i)$$

divise m_f.
Si x_i est un vecteur propre associé à λ_i, alors il est clair que $(f - \lambda_i \mathrm{Id}_E)(x_i) = 0$ et donc

$$M(f)(x_i) = \left(\prod_{j\neq i}(\lambda_j \mathrm{Id}_E - f)\right)(\lambda_i \mathrm{Id}_E - f)(x_i) = 0.$$

Si f est diagonalisable, alors en vertu du théorème 12 $E = \bigoplus_{i=1}^{r} E_{\lambda_i}$. Donc tout $x \in E$ se décompose en $x = x_1 + \cdots + x_r$ avec $x_i \in E_{\lambda_i}$, puis d'après le résultat précédent

$$M(f)(x) = \sum_{i=1}^{r} M(f)(x_i) = 0.$$

On obtient ainsi (ii).
$(ii) \Rightarrow (iii) \Rightarrow (iv)$: trivial.
$(iv) \Rightarrow (i)$: soit P un polynôme annulateur de f scindé à racines simples, que l'on peut toujours supposer unitaire (quitte à le multiplier par une constante) :

$$P(X) = \prod_{i=1}^{s}(X - \mu_i) \quad \text{avec les } \mu_i \text{ deux à deux distincts.}$$

Comme les polynômes $X - \mu_i$ sont premiers entre eux on peut appliquer le corollaire 26, ce qui donne

$$E = \mathrm{Ker}[P(f)] = \bigoplus_{i=1}^{s} \mathrm{Ker}(f - \mu_i \mathrm{Id}_E).$$

Si dans cette décomposition, on ne garde que les μ_i tels que $\mathrm{Ker}(f - \mu_i \mathrm{Id}_E)$ ne soit pas réduit à $\{0\}$, on a obtenu une décomposition de E en somme de sous-espaces propres de f. f est donc diagonalisable. □

Corollaire 40. *Soit $f \in \mathcal{L}(E)$ diagonalisable. Soit F un sous-espace de E, stable par f. Alors* l'endomorphisme induit *f_F est diagonalisable.*

Démonstration. Si f est diagonalisable, alors d'après le résultat précédent, il existe $P \in K[X]$ scindé à racines simples tel que $P(f) = 0$. On a bien sur aussi $P(f_F) = 0$. Donc à nouveau d'après le théorème 27, f_F est diagonalisable. □

Corollaire 41. *Soient f et g deux endomorphismes diagonalisables de E qui commutent. Alors il existe une base de diagonalisation commune pour f et g.*

De même, si A et B sont deux matrices diagonalisables qui commutent, alors il existe une matrice inversible P et deux matrices diagonales D et D' telles que

$$A = PDP^{-1} \quad B = PD'P^{-1}$$

Démonstration. Si f est diagonalisable, alors on a $E = \oplus_{i=1}^{r} E_{\lambda_i}$. Comme f et g commutent il est facile de voir que chaque E_{λ_i} est stable par g. L'endomorphisme g_i, induit par g sur E_{λ_i} est d'après le résultat précédent diagonalisable. On peut donc trouver une base $(e_1^i, \ldots, e_{\gamma_i}^i)$ de E_{λ_i} qui est une base de vecteurs propres pour g_i. Bien sûr, chaque e_k^i est un vecteur propre de f associé à la valeur propre λ_i.
En rassemblant les bases obtenues pour chaque E_{λ_i}, on obtient une base de E, qui est une base commune de vecteurs propres pour f et g. □

Ce résultat peut s'étendre sans trop de difficultés au cas de p endomorphismes diagonalisables qui commutent.

VI. Sous-espaces caractéristiques

Dans cette partie on suppose que f est un endomorphisme scindé d'un espace vectoriel E de dimension n. On notera $\{\lambda_1, \ldots \lambda_r\}$ l'ensemble de ses valeurs propres distinctes. Son polynôme caractéristique s'écrit donc

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i},$$

où α_i est l'ordre de multiplicité de la valeur propre λ_i. De plus, par le Théorème de Cayley-Hamilton son polynôme minimal s'écrit

$$m_f(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\beta_i},$$

avec pour $1 \leq i \leq r$, $1 \leq \beta_i \leq \alpha_i$.

VI.1. Définition et premières propriétés.

Définition 42. *On reprend les notations ci-dessus. On appelle* sous-espace caractéristique *de f relatif à la valeur propre λ_i le sous-espace vectoriel de E défini par $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$, où α_i est l'ordre de multiplicité de la valeur propre λ_i.*

Remarque 43. *Si $\alpha_i = 1$ alors $N_{\lambda_i} = E_{\lambda_i}$, autrement dit pour une valeur propre simple le sous-espace caractéristique est égal au sous-espace propre.*

Exemple 44. *Soient E un espace vectoriel de dimension 3 et f l'endomorphisme de E de matrice*

$$M = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

dans une base (e_1, e_2, e_3) de E. Son polynôme caractéristique est $\chi_f(X) = (2 - X)^2(3 - X)$, de sorte que son polynôme minimal est $m_f(X) = (X - 2)^k(X - 3)$ avec $k = 1$ ou $k = 2$. Par le calcul on vérifie que la matrice $(M - 2\mathrm{I}_3)(M - 3I_3)$ n'est pas nulle, donc $m_f(X) = (X - 2)^2(X - 3)$. Comme les racines de m_f ne sont pas toutes simples, f n'est pas diagonalisable. Comme 3 est une valeur propre simple, E_3 est un sous-espace de dimension 1, et d'après la forme de la matrice on peut dire qu'il est engendré par e_3. L'endomorphisme f n'étant pas diagonalisable, la dimension de E_3 nous permet d'affirmer que E_2 est de dimension 1, et d'après M qu'il est engendré par e_1. Pour déterminer N_2 il faut résoudre $(M - 2\mathrm{I}_3)^2 v = 0$ et on trouve que c'est le plan engendré par e_1 et e_2. D'autre part, on a $N_3 = E_3$, car 3 est une valeur propre simple. On peut remarquer que (e_1, e_2, e_3) est une base de E, et par conséquent que $E = N_2 \oplus N_3$. Ce résultat va être montré de manière générale dans la suite.

Proposition 45. *Le sous-espace propre E_{λ_i} associé à la valeur propre λ_i est inclus dans le sous-espace caractéristique N_{λ_i}.*

Démonstration. Soit $x \in E_{\lambda_i}$, c'est-à-dire tel que $(f - \lambda_i \mathrm{Id}_E)(x) = 0$. On a

$$(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}(x) = (f - \lambda_i \mathrm{Id}_E)^{\alpha_i - 1}\big((f - \lambda_i \mathrm{Id}_E)(x)\big) = 0.$$

Ainsi x appartient à N_{λ_i}. □

En particulier ceci entraîne que la dimension de N_{λ_i} est toujours supérieure ou égale à 1, car celle de E_{λ_i} l'est.

Proposition 46. *Les sous-espaces caractéristiques sont stables par f.*

Démonstration. Soit $x \in N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big]$, c'est-à-dire tel que $\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big](x) = 0$. On a alors

$$\big[(f - \lambda_i \mathrm{Id}_E)_i^{\alpha}\big] f(x) = f\big(\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big](x)\big) = 0.$$

Donc $f(x) \in N_{\lambda_i}$. □

Théorème 47.

On a

$$E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}.$$

Démonstration. On a

$$\chi_f(X) = (-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}.$$

D'après le Théorème de Cayley-Hamilton, $\chi_f(f) = 0$, c'est-à-dire que $E = \mathrm{Ker}\big[\chi_f(f)\big]$. De plus si $i \neq j$, alors $(X - \lambda_i)^{\alpha_i}$ et $(X - \lambda_j)^{\alpha_j}$ sont des polynômes premiers entre eux. Donc le corollaire 26 entraîne que

$$\mathrm{Ker}\big[\chi_f(f)\big] = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r},$$

et donc le résultat. □

Proposition 48. *Pour tout $1 \leq i \leq r$, on a $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$, où β_i est l'ordre de multiplicité de la racine λ_i dans le polynôme minimal.*

Démonstration. On observe les faits suivants :

— Comme $\beta_i \leq \alpha_i$ la même démonstration que pour la proposition 31 montre que :

$$\mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big] \subset \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big].$$

— On a $m_f(f) = 0$, c'est-à-dire $E = \mathrm{Ker}\big[m_f(f)\big]$. De plus, si $i \neq j$, alors $(X - \lambda_i)^{\beta_i}$ et $(X - \lambda_j)^{\beta_j}$ sont des polynômes premiers entre eux. Donc par le corollaire 26 on obtient

$$E = \mathrm{Ker}\big[m_f(f)\big] = \mathrm{Ker}\big[(f - \lambda_1 \mathrm{Id}_E)^{\beta_1}\big] \oplus \cdots \oplus \mathrm{Ker}\big[(f - \lambda_r \mathrm{Id}_E)^{\beta_r}\big].$$

— La même démarche appliquée au polynôme caractéristique montre que

$$E = \mathrm{Ker}\big[\chi_f(f)\big] = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}.$$

En combinant ceci on déduit que $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$ pour tout $1 \leq i \leq r$. □

Théorème 49.

f est diagonalisable si et seulement si les deux conditions suivantes sont vérifiées :
(i) f est scindé.
(ii) Pour toute valeur propre λ_i de f on a $N_{\lambda_i} = E_{\lambda_i}$.

Démonstration. On a déjà vu que si f est diagonalisable alors f est scindé et de plus on a

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}.$$

D'autre part, f étant scindé, d'après le théorème 34 on a aussi $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$ et d'après la proposition 31, $E_{\lambda_i} \subset N_{\lambda_i}$. Tout ceci entraîne que $E_{\lambda_i} = N_{\lambda_i}$.

Réciproquement, si f est scindé d'après le théorème 34, $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$. Ainsi comme $E_{\lambda_i} = N_{\lambda_i}$, $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$, c'est-à-dire que f est diagonalisable. □

VI.2. Applications linéaires restreintes.

Proposition 50. *La restriction $f_i = f|_{N_{\lambda_i}}$ de f au sous-espace caractéristique N_{λ_i} est un endomorphisme de N_{λ_i}. Les polynômes minimal et caractéristique de f_i sont*

$$m_{f_i}(X) = (X - \lambda_i)^{\beta_i} \text{ et } \chi_{f_i}(X) = (-1)^{\alpha_i}(X - \lambda_i)^{\alpha_i}.$$

De plus on a $\dim(N_{\lambda_i}) = \alpha_i$.

Démonstration. La restriction de f à N_{λ_i} est une application linéaire de N_{λ_i} dans $f\big(N_{\lambda_i}\big)$. Or d'après la proposition 33, N_{λ_i} est stable par f et donc par sa restriction. On en déduit que $f_i \in \mathcal{L}\big(N_{\lambda_i}\big)$.

On sait que

$$N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\alpha_i}\big] = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big].$$

On en déduit donc $(f_i - \lambda_i \mathrm{Id}_{N_{\lambda_i}})^{\beta_i} = 0$, c'est-à-dire que le polynôme minimal de f_i divise le polynôme $(X - \lambda_i)^{\beta_i}$. Supposons $\deg(m_{f_i}) = \gamma_i < \beta_i$, alors le polynôme

$$Q(X) = (X - \lambda_i)^{\gamma_i} \prod_{j \neq i} (X - \lambda_j)^{\beta_j}$$

s'annule en f et

$$\gamma_i + \sum_{j \neq i} \beta_j < \sum_{j=1}^{r} \beta_j = \deg(m_f).$$

Ceci est impossible car le polynôme minimal de f est le polynôme unitaire de plus bas degré vérifiant cette propriété. Ainsi on a $\gamma_i = \beta_i$ et $m_{f_i}(X) = (X - \lambda_i)^{\beta_i}$.

Supposons que $\dim(N_{\lambda_i}) = \gamma_i$. Comme E est la somme directe des N_{λ_i} il existe une base $\mathcal{B}$ de E telle que $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, où $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\gamma_i,i})$ est une base de N_{λ_i}. Dans cette base la matrice de f s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}$$

où les M_i sont les matrices des f_i dans les bases $\mathcal{B}_i$. Or le polynôme minimal de f_i est $m_{f_i}(X) = (X-\lambda_i)^{\beta_i}$. Donc la seule valeur propre de M_i (c'est-à-dire de f_i) est λ_i et son polynôme caractéristique est égal à $(-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i}$. On obtient donc l'égalité suivante :

$$\begin{aligned}(-1)^n\prod_{i=1}^r(X-\lambda_i)^{\alpha_i} &= \chi_f(X) = \prod_{i=1}^r \det(M_i - X\mathrm{I}_{\gamma_i}) = \prod_{i=1}^r \chi_{f_i}(X) = \prod_{i=1}^r (-1)^{\gamma_i}(X-\lambda_i)^{\gamma_i} \\ &= (-1)^n\prod_{i=1}^r(X-\lambda_i)^{\gamma_i}\end{aligned}$$

car

$$\sum_{i=1}^r \gamma_i = \sum_{i=1}^r \dim(N_{\lambda_i}) = \dim(E) = n.$$

On en déduit que $\gamma_i = \alpha_i$, c'est-à-dire que $\dim(N_{\lambda_i}) = \alpha_i$. De plus on obtient que $\chi_{f_i}(X) = (-1)^{\alpha_i}(X-\lambda_i)^{\alpha_i}$. □

VI.3. Trigonalisation des matrices en blocs relatifs aux sous-espaces caractéristiques.

On a vu que si on prend $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ avec $\mathcal{B}_i = (e_{1,i}, \ldots, e_{\alpha_i,i})$ une base de N_{λ_i}, alors la matrice de f dans $\mathcal{B}$ s'écrit

$$M = \begin{pmatrix} M_1 & & & & \\ & M_2 & & & \\ & & \ddots & & \\ & & & M_{r-1} & \\ & & & & M_r \end{pmatrix}.$$

Pour trigonaliser M il suffit alors de trigonaliser chaque M_i.

Exemple : On veut trigonaliser l'endomorphisme f dont la matrice dans la base canonique est

$$\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}.$$

On commence par chercher le polynôme caractéristique de f. On obtient

$$\chi_f(X) = (X-1)^2(X+1)^2.$$

f a donc deux valeurs propres doubles, 1 et -1. On détermine alors les sous-espaces caractéristiques. Pour $\lambda = -1$, on obtient :

$$N_{-1} = \mathrm{Ker}[(f + 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M + \mathrm{I}_4)^2].$$

On a donc

$$(M+\mathrm{I}_4)^2\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} 12z - 8t &= 0 \\ 8z - 4t &= 0 \\ 12z - 8t &= 0 \\ 8z - 4t &= 0 \end{cases} \iff \begin{cases} 3z &= 4t \\ 2z &= t \end{cases}$$

Une base de N_{-1} est donc donnée par u_1 et u_2 de coordonnées dans la base canonique

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Pour $\lambda = 1$, $N_1 = \mathrm{Ker}[(f - 1\mathrm{Id}_E)^2] = \mathrm{Ker}[(M - \mathrm{I}_4)^2]$. On a donc

$$(M-\mathrm{I}_4)^2\begin{pmatrix}x\\y\\z\\t\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\iff\begin{cases}-12x+16y+12z-16t=0\\-16x+20y+16z-20t=0\end{cases}\iff\begin{cases}x=z\\y=t\end{cases}.$$

Une base de N_1 est donc donnée par u_3 et u_4 de coordonnées dans la base canonique $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$ et $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$.

La matrice de passage de la base canonique à la nouvelle base (u_1,u_2,u_3,u_4), ainsi que son inverse, sont données par

$$P=\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{pmatrix}\quad\text{et}\quad P^{-1}=\begin{pmatrix}1&0&-1&0\\0&1&0&-1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Donc la matrice de f dans la nouvelle base s'écrit :

$$\widetilde{M}=P^{-1}MP=\begin{pmatrix}3&-4&0&0\\4&-5&0&0\\0&0&3&-2\\0&0&2&1\end{pmatrix}.$$

On trigonalise alors chacune des sous matrices

$$M_1=\begin{pmatrix}3&-4\\4&-5\end{pmatrix}\quad\text{et}\quad M_2=\begin{pmatrix}3&-2\\2&1\end{pmatrix}.$$

Le polynôme caractéristique de M_1 est égal à $(X+1)^2$. Le sous-espace propre de M_1 associé à la valeur propre -1 est déterminé par

$$(M_1+\mathrm{I}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\iff\begin{cases}4x_1-4x_2&=&0\\4x_1-4x_2&=&0\end{cases}\iff x_1=x_2.$$

Une base de cet espace est donc $\begin{pmatrix}1\\1\end{pmatrix}$ dans (u_1,u_2), i.e.

$$v_1=u_1+u_2=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}.$$

On complète par un second vecteur de N_{-1} non colinéaire à v_1, par exemple $v_2=u_1$.
Le polynôme caractéristique de M_2 est égal à $(X-1)^2$. Le sous-espace propre de M_2 associé à la valeur propre 1 est déterminé par

$$(M_2-\mathrm{I}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\iff\begin{cases}2x_3-2x_4&=&0\\2x_3-2x_4&=&0\end{cases}\iff\{\ x_3\ =\ x_4$$

Une base de cet espace est donc $\begin{pmatrix}1\\1\end{pmatrix}$ dans (u_3,u_4), i.e.

$$v_3=u_3+u_4=\begin{pmatrix}0\\0\\1\\1\end{pmatrix}.$$

On complète par un second vecteur de N_1 non colinéaire à v_3, par exemple $v_4 = u_3$.
La matrice de passage de la base (u_1, u_2, u_3, u_4) à la base (v_1, v_2, v_3, v_4) et son inverse, sont égales à

$$P' = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad P'^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Donc la matrice de f dans la base (v_1, v_2, v_3, v_4) s'écrit :

$$M' = (P')^{-1}\widetilde{M}(P') = \begin{pmatrix} -1 & 4 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

VII. Endomorphismes nilpotents

VII.1. Caractérisation des endomorphismes nilpotents.

Définition 51. *On dit qu'un endomorphisme f (respectivement une matrice M) est* nilpotent *(respectivement* nilpotente*) si et seulement si il existe un entier $k \geq 1$ tel que $f^k = 0$, l'endomorphisme identiquement nul, (respectivement $M^k = 0$, la matrice nulle). Le plus petit entier $k \geq 1$ vérifiant ceci, est alors appelé l'*indice de nilpotence *de f (respectivement de M).*

Proposition 52. *Un endomorphisme d'un espace vectoriel de dimension n est nilpotent si et seulement si son polynôme caractéristique est $(-1)^n X^n$.*

Démonstration. Si f est un endomorphisme nilpotent, alors il existe un entier $k \geq 1$ tel que $f^k = 0$, c'est-à-dire que le polynôme X^k appartient à l'idéal annulateur de f. Donc par définition, le polynôme minimal de f divise X^k ; autrement dit il est de la forme X^j avec j un entier tel que $1 \leq j \leq k$. L'unique racine de ce polynôme est donc 0, et donc l'unique valeur propre de f est 0. Comme l'espace vectoriel est de dimension n, le polynôme caractéristique de f est donc $(-1)^n X^n$. Réciproquement, si le polynôme caractéristique de f est $(-1)^n X^n$, d'après le Théorème de Cayley-Hamilton $\chi_f(f) = f^n = 0$, c'est-à-dire que f est nilpotent. □

Une conséquence de ce résultat est que le polynôme minimal d'un endomorphisme nilpotent est égal à X^p avec $p \leq n$ (ceci montre au passage que l'indice de nilpotence est inférieur ou égal à la dimension de E). D'autre part, comme il n'a qu'une seule valeur propre qui est zéro, il n'est pas diagonalisable sauf s'il est nul. Mais il est trigonalisable.

VII.2. Décompositions de Dunford et de Jordan.

Théorème 53.

Soit f un endomorphisme d'un espace vectoriel E. Si f est scindé alors il existe un unique couple (u, v) d'endomorphismes tel que

- *$f = u + v$.*
- *u est diagonalisable et v est nilpotent.*
- *u et v commutent.*

Démonstration. Montrons que le couple (u, v) existe. Comme f est scindé d'après le théorème 34 on a $E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_r}$, où les $N_{\lambda_i} = \mathrm{Ker}\big[(f - \lambda_i \mathrm{Id}_E)^{\beta_i}\big]$ sont les sous-espaces caractéristiques de f correspondant aux différentes valeurs propres $(\lambda_i)_{1 \leq i \leq r}$ de f. On définit alors u et v par leurs restrictions aux sous-espaces caractéristiques de la manière suivante :

$$u_{|N_{\lambda_i}} = \lambda_i \mathrm{Id}_{N_{\lambda_i}} \text{ et } v_{|N_{\lambda_i}} = f_{|N_{\lambda_i}} - \lambda_i \mathrm{Id}_{N_{\lambda_i}}.$$

Comme les sous-espaces caractéristiques sont supplémentaires et que chaque restriction de u est diagonale, u l'est aussi. D'autre part, par définition chaque restriction de v est nilpotente d'ordre $n_i \leq \beta_i$ et donc v sera aussi nilpotente d'ordre $\max_{1\leq i\leq r}(n_i)$. Enfin il est clair que pour tout $1 \leq i \leq r$, $u_{|N_{\lambda_i}}$ commute avec $v_{|N_{\lambda_i}}$ car l'identité commute avec tout endomorphisme, donc u et v commutent. Nous admettons l'unicité de la décomposition. □

Exemple 54. *Soit la matrice*

$$M = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}.$$

On vérifie par le calcul que son polynôme caractéristique est $\chi_M(X) = -X^2(X+1)$. Les valeurs propres sont donc : -1 valeur propre simple et 0 valeur propre double. Le sous-espace caractéristique associé à la valeur propre -1, est égal au sous-espace propre associé à cette même valeur propre. Pour le déterminer on résout $(M+\mathrm{I}_3)v = 0$. On trouve qu'il est engendré par $e_1 = (1,-1,-2)$. Pour le sous-espace caractéristique associé à la valeur propre 0 on résout $(M-0.\mathrm{I}_3)^2 v = M^2 v = 0$ et on constate qu'il est engendré par $e_2 = (1,2,0)$ et $e_2 = (0,1,1)$. Nous allons maintenant écrire la matrice dans la nouvelle base. La matrice de passage de la base canonique de $\mathbb{R}^3$ à la base (e_1,e_2,e_3) est

$$P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix} \text{ et son inverse } P^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 4 & -2 & 3 \end{pmatrix}.$$

On en déduit que

$$\widetilde{M} = P^{-1}MP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \widetilde{D} + \widetilde{N}.$$

La décomposition de Dunford de M est donc $M = D + N$ avec

$$D = P\widetilde{D}P^{-1} = \begin{pmatrix} -2 & 1 & -1 \\ 2 & -1 & 1 \\ 4 & -2 & 2 \end{pmatrix} \quad \text{et} \quad N = P\widetilde{N}P^{-1} = M - D = \begin{pmatrix} 4 & -2 & 3 \\ 8 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}.$$

On peut vérifier que $DN = ND$.

Venons-en à la écomposition de Jordan.

Définition 55. *On appelle* bloc de Jordan *de taille p et de valeur propre λ une matrice carrée d'ordre p de la forme*

$$J_{\lambda,p} = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \ldots & 0 & \lambda \end{pmatrix}.$$

Théorème 56.

Soit f un endomorphisme scindé de E. On suppose que le polynôme caractéristique de f est

$$(-1)^n \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i}, \qquad \text{avec } \sum_{i=1}^{r} \alpha_i = n,$$

où les λ_i, $1 \leq i \leq r$ sont deux à deux distincts. Il existe alors une base de E dans laquelle la matrice de f est diagonale par blocs de la forme :

$$\begin{pmatrix} J_{\lambda_1,n_{1,1}} & 0 & \cdots & 0 \\ 0 & J_{\lambda_1,n_{1,2}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & J_{\lambda_r,n_{r,s_r}} \end{pmatrix},$$

où pour $1 \leq i \leq r$ et $1 \leq j \leq s_i$, le bloc diagonal $J_{\lambda_i,n_{i,j}}$ est un bloc de Jordan de taille $n_{i,j}$ de valeur propre λ_i. Pour $1 \leq i \leq r$, on a $\sum_{j=1}^{s_i} n_{i,j} = \alpha_i$, et on impose $n_{i,1} \geq n_{i,2} \geq \cdots \geq n_{i,s_i}$. L'entier $n_{i,1}$ est égal à la multiplicité de λ_i comme racine du polynôme minimal.

On peut remarquer que la matrice de Jordan donnée par le théorème est triangulaire supérieure puisque les blocs de Jordan le sont. Sur la diagonale figurent les valeurs propres, avec la multiplicité qu'elles ont dans le polynôme caractéristique, et les coefficients de la diagonale secondaire $\{(i,j), j = i+1\}$ sont égaux à 1 ou 0. Tous les autres coefficients sont nuls.
Les valeurs propres étant données et indexées, la matrice de Jordan est entièrement déterminée par la donnée des entiers $n_{i,1},\ n_{i,2},\ \ldots,\ n_{i,s_i}$. Avec ces hypothèses, elle est unique et est appelée *forme normale de Jordan*.